RICE UNIVERSITY

# Engineering *Escherichia coli* for the production of polyketide-based platform chemicals

by

**John Park**

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

**Master of Science**

APPROVED, THESIS COMMITTEE

Ramon Gonzalez, Chair
William W. Akers Assistant Professor in
Chemical and Biomolecular Engineering

Laura Segatori
T. N. Law Assistant Professor in Chemical
and Biomolecular Engineering

Sibani Lisa Biswal
Assistant Professor in Chemical and
Biomolecular Engineering

HOUSTON, TEXAS
April 2011

# ABSTRACT

**Engineering *Escherichia coli* for the production of polyketide-based platform chemicals**

by

John Park

The current chemical industry produces a diverse array of industrial chemicals from a handful of highly reduced byproducts (termed "platform chemicals") derived from oil refining. However, petroleum is a non-renewable resource, and increases in its cost have created pressure to convert the chemical industry into one that is renewable to ensure its long-term viability. To complete this objective, one approach is the conversion of biomass to platform chemicals through fermentation by *Escherichia coli*. One such platform chemical is methyl ketone, which can be readily converted to dienes that can directly replace existing platform chemicals such as ethylene. To bestow non-native methyl ketone production capability to *E. coli* from glucose, the polyketide biosynthesis pathway was exploited in conjunction with grafting in a heterologous methyl ketone synthesis pathway found in wild tomato species *Solanum habrochaites* to produce the methyl ketones. Cultivation under microaerobic conditions improved titers and yields, and further engineering to knock out the native competitive pathways that become activated under microaerobic conditions led to significantly improved strains. The final strain, *ΔadheΔldhaΔptaΔpoxB* [pTrcHis2A-shmks2-mks1], produced up to 450 mg/L of methyl ketones at 17 mg of methyl ketones produced per gram of glucose consumed under optimized operating conditions in minimal media supplemented with glucose.

# ACKNOWLEDGEMENTS

I would like to thank the many people who have guided me along the way in assembling this thesis. First, I am immensely grateful to have an advisor such as Dr. Ramon Gonzalez, for all the invaluable guidance, support, instruction, and in general putting up with me during my research. Next, I would like to thank Dr. Ka-Yiu San for his advice and suggestions over the past two years, and Dr. Laura Segatori and Dr. Lisa Biswal for sitting in my committee. Finally, I'd like to thank the rest of the research group and collaborators: I would like to thank Dr. Mai Li from the San Lab for helping me run GC-MS for this study, and for providing the initial MKS plasmids. I would like to thank Dr. Matthew Blankschien for providing me with knockout strains, guiding me through various molecular biology techniques and strain construction, and helping me set up the fermentors. I would like to thank James Clomburg for everyday advice and help with the fermentors. I would like to thank the BioNMR facility at Rice, in particular Dr. Sean Moran for assistance with the NMR. I would like to thank Dr. Eran Pichersky and the rest of the Center for Biorenewable Chemicals (CBiRC) for the highly informative seminars, meetings, and general knowledge-sharing that helped move this study along.

# TABLE OF CONTENTS

# List of Tables and Figures

# 1. INTRODUCTION

## 1.1. THE CHEMICAL INDUSTRY AND PETROLUEM

Since the discovery of fire, humans have been harnessing bio-based products and inorganic materials as energy sources and feedstocks for the production of materials. However, starting with coal during the Industrial Revolution, society has increasingly depended on non-renewable fossil carbon for the manufacture of these materials. Over time, this dependence has expanded into other types of fossil carbon which have historically been inexpensive and abundant such as petroleum and natural gas. Currently, ninety percent of the world's commodity chemicals are derived from these fossil carbon sources, with only a small remainder derived from bio-renewable carbon sources. However, in the last 10 years, a number of developments have restored interest in using renewable sources of carbon. Primarily, rapid increases in the global demand for energy have dramatically raised the consumption rate and cost of fossil fuels, which has led to some concern over the remaining supply of fossil fuels and consequently the future of the chemical industry. These developments have spurred advancements in oil discovery and recovery efforts, which have been moderately successful in increasing the supply of oil albeit, in some cases, at the cost of negative environmental impacts. Spirited research into alternative energy sources such as solar, wind, and nuclear have yielded promising results, moving us closer and closer to a renewable society. However, the fact remains that these can only be short-term solutions for the chemical industry. In addition to fossil carbon being inherently non-renewable, the production of industrial chemicals relies on a carbon backbone, and thus it is imperative that a way is found for their production from a renewable carbon source.

The chemical sector is the largest industrial consumer of energy products, consuming approximately 15 percent of the world's total delivered energy and making up 29 percent of the world's industrial energy consumption in 2006 as shown in Figure 1 [1]. It is important to note that within the chemical sector, the majority of the energy consumed is used as feedstock for the production of industrial organic chemicals, as opposed to its traditional use as a fuel source. Though the rest of the world's energy is consumed through the transportation sector, the value of all chemicals derived from energy products is approximately equal to that of transportation fuels, illustrating the importance and large economic impact of the petrochemical industry. Indeed, the products derived from industrial chemicals have significantly impacted society through their myriad applications ranging from food additives to plastics. Additionally, these products see widespread use due to their low unit costs, made possible by highly optimized processing of hydrocarbons derived from fossil carbon. Understanding this tightly-knit association of the chemical industry with petroleum and gas is the first key step in assessing viable approaches to a sustainable solution for the chemical industry.

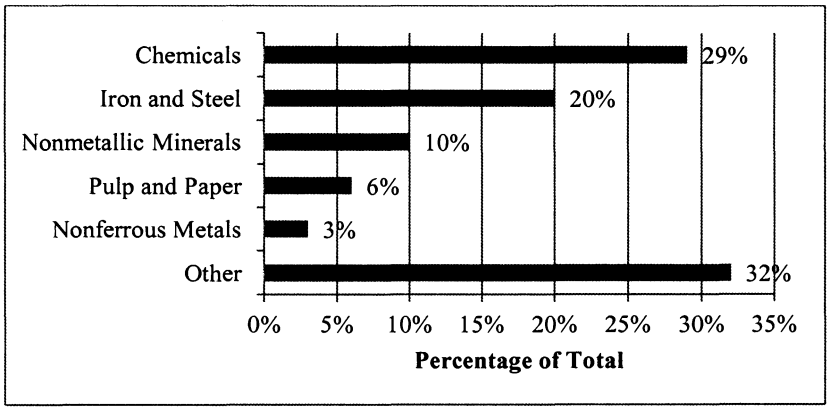



**Figure 1. World industrial sector energy consumption**

The heavy dependence of the chemical industry on oil and natural gas is clear when one looks at its inception. The chemical industry as it is known today was formed around the need to use hydrocarbon by-products generated by the refining process of crude oil that were too volatile for use as transportation fuels. These undergo fluid catalytic cracking to form ethylene and propylene, which along with benzene are converted through various chemical processes such as oxidation, alkylation, and polymerization to a vast array of chemical products [2, 3]. For example, Figure 2 shows a number of large-volume industrial chemicals that are derived from propylene and ethylene.

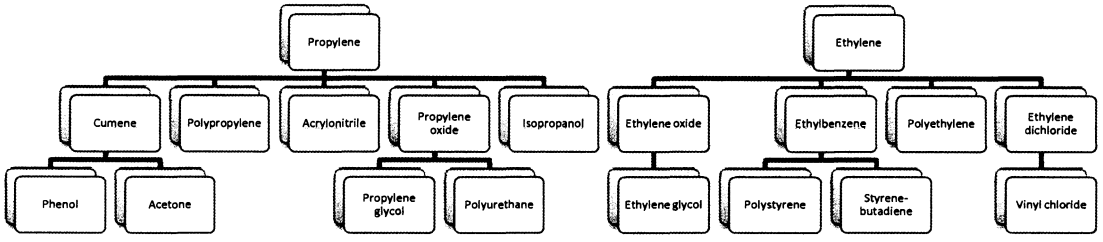



**Figure 2. Large-volume chemicals derived from two example platform chemicals**

(Source: Center for Biorenewable Chemicals (CBiRC), http://www.cbirc.iastate.edu/.)

Over the past 80 years, improvements in feedstock conversion and downstream chemical processes have led to highly optimized utilization of oil and gas by-products in the industry. Central to this efficiency is the use of the aforementioned building blocks, known as platform chemicals, that are derived from these fossil carbon sources. This is best illustrated in Table 1, which lists the top 20 industrial chemicals used in the United

States [4]. Platform chemicals are at or near the top of the list, confirming their significance in the industry.

**Table 1. List of the top 20 industrial chemicals used in the United States**

| | | | |
|---|---|---|---|
| 1 | **Ethylene** | 11 | Ethylene oxide |
| 2 | **Propylene** | 12 | *p*-xylene |
| 3 | Ethylene dichloride | 13 | Cumene |
| 4 | Methanol | 14 | Ethylene glycol |
| 5 | Vinyl chloride | 15 | Butadiene |
| 6 | Benzene | 16 | Phenol |
| 7 | Ethylbenzene | 17 | Acetic acid |
| 8 | Styrene | 18 | Acrylonitrile |
| 9 | Terephthalic acid | 19 | α-olefins |
| 10 | Formaldehyde | 20 | Propylene oxide |

The reliance of the current chemical industry on fossil carbon-derived platform chemicals means it is vulnerable to the cost and supply of oil and natural gas. On top of the inherent price volatility of these resources, the average prices are forecasted to increase in the near future according to the most recent International Energy Outlook document published by the U.S. Department of Energy as shown in Figure 3. These two economic trends will negatively impact the growth and investment of the chemical industry [5].

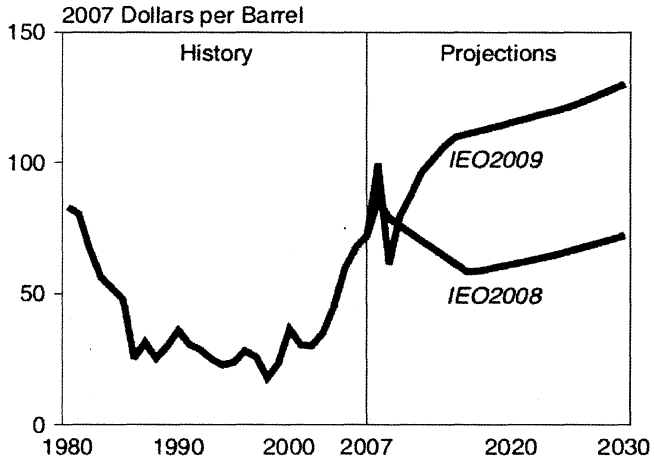



**Figure 3. World oil prices in the 2008 and 2009 International Energy Outlook (IEO) reference cases**

(Source: Energy Information Administration (EIA), Annual Energy Outlook 2009, DOE/EIA-0383(2009) (Washington, DC, March 2009), web site www.eia.doe.gov/oiaf/aeo.)

More importantly, however, is the dwindling supply of fossil carbon resources. Producing carbon-based chemicals from fossilized carbon is inherently non-sustainable, as it requires millions of years to form naturally, a rate which is far surpassed by the rate of its current consumption, which is predicted to continually rise as shown in Figure 4. Along with rising consumption, the majority of forecasts of world oil production predict the maximum to occur between 2010 and 2020, followed by a rapid decline through the following decades as shown in Figure 5. Eventually, the supply of fossil carbon will either be exhausted or prohibitively costly to acquire, so a long-term solution is required for the transformation of the current petroleum-based chemical industry to one that will be able to use a feedstock with a timescale of formation equal to that of its utilization.

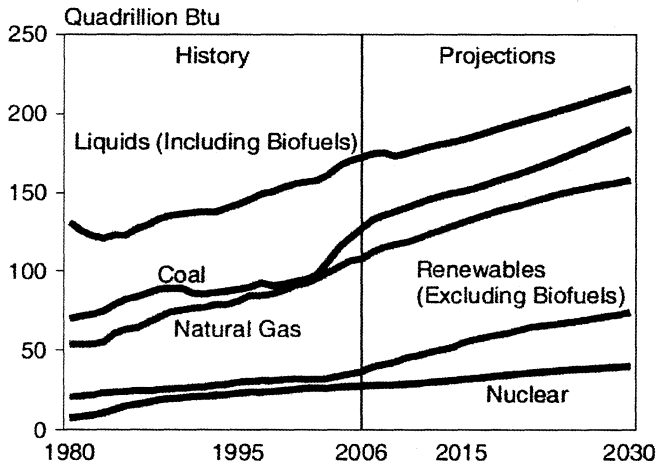



**Figure 4. World energy consumption by fuel type**

(Sources: 2006: Energy Information Administration (EIA), International Energy Annual 2006 (June-December 2008), web site www.eia.doe.gov/iea. Projections: EIA, World Energy Projections Plus (2009). )

Biomass, with its short formation time, low cost, and ability to supply carbon, is currently the only option as a renewable feedstock for the chemical industry [6]. The impetus to switch to this option is, as aforementioned, dependent on the remaining supply of fossil carbon. However, unlike its use as an energy source, which could be resolved through the use of alternative energy technologies such as solar, wind, and nuclear, the production of organic chemicals can only attain sustainability through the use of biomass, given the inherent need for a carbon backbone in industrial chemicals.

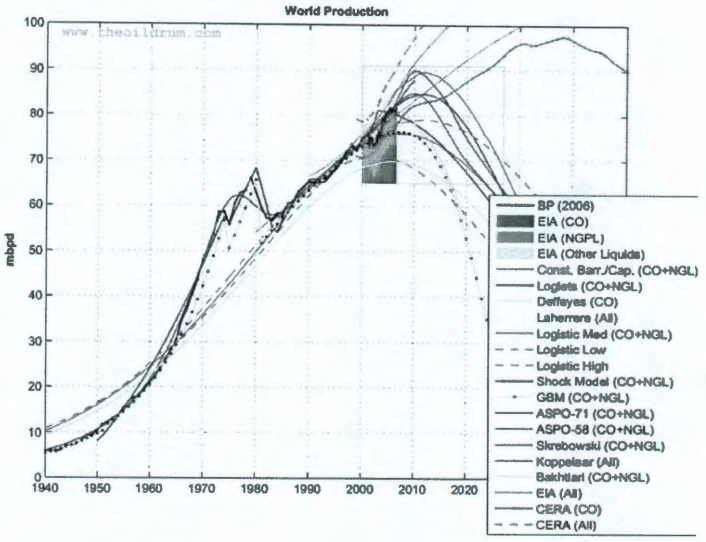



**Figure 5. World oil production (crude oil and natural gas) along with various forecasting models using historical data between 1940 to 2006 from the Energy Information Administration (DOE)**

(source: Foucher, S. Production Forecasts and EIA Oil Production Numbers.http://www.theoildrum.com/story/2006/11/13/225447/79)

## 1.2. INTEGRATION OF A BIO-BASED CHEMICAL INDUSTRY WITHIN A BIO-REFINERY FRAMEWORK

In order to move the chemical industry toward biomass without significant costs, processes will need to be developed that will allow efficient conversion of feedstocks. Mirroring the petrochemical industry, which is highly integrated and productive, many believe a bio-based chemical industry will be successful if built upon a platform chemical approach [7, 8]. Werpy and Peterson (2004), through a collection of expert opinion from industry, introduced a list of 12 potential biologically-derived platform chemicals as shown in Figure 6, which could be converted to a number of chemical products such as acrylic acid, 1,3-propanediol, acrylamide, and acrylonitrile. However, the proposed synthetic routes, either through a biological process or the chemical conversion of a biological precursor, are currently not economically viable due to their low efficiency,

but further research is warranted to improve these processes [6]. Since the listed chemicals are able to be produced through the fermentation of sugars in biomass, one potential way to increase the viability of these processes is to co-locate them at a site which also produces biofuels, a concept which is called a bio-refinery [9].

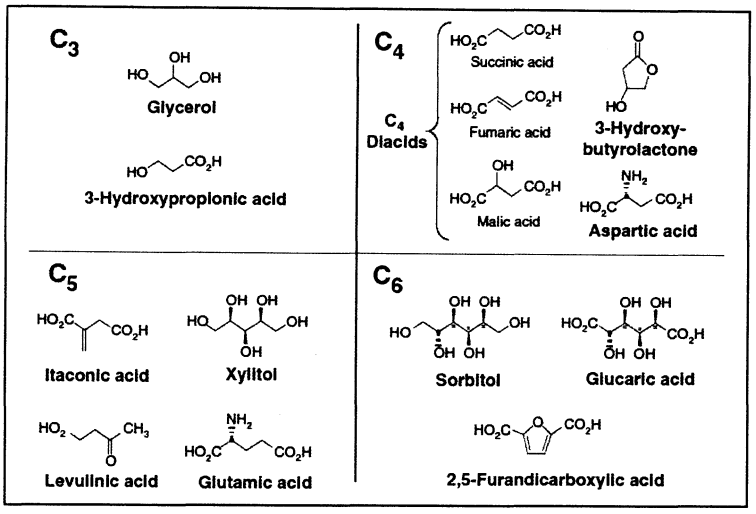



**Figure 6. A list of 12 potential bio-based platform chemicals as presented in Werpy and Peterson (2004) [8]**

A bio-refinery is analogous to a crude oil refinery, wherein multiple fuel and platform chemicals are co-produced from a single feedstock at a single location. The difference lies in the feedstock that is utilized; in a traditional refinery, crude oil is the main feedstock, whereas in a bio-refinery, biomass is the main feedstock. By using both a portion of the feedstock as well as by-products from the production of biofuel, chemicals can be made to offset costs and maximize the value of the biomass feedstock. One current example of a bio-refinery is the corn wet mill, which produces ethanol along with proteins, lactic acid, maltodextrins, and smaller specialty chemicals [10, 11]. The scheme for

this process is in Figure 7, which illustrates how one feedstock can lead to a diverse array of products. The co-production of chemicals and biofuels at a single site vastly increases the economic viability of the production of both biofuels and bio-based chemicals through integration of feedstock and by-product handling, economy of scale, internal management of energy requirements, and sharing of utilities [12]. However, before such a concept can be implemented, the processes involved with the production of chemicals from both the biomass feedstock and biofuel waste streams need to be developed.

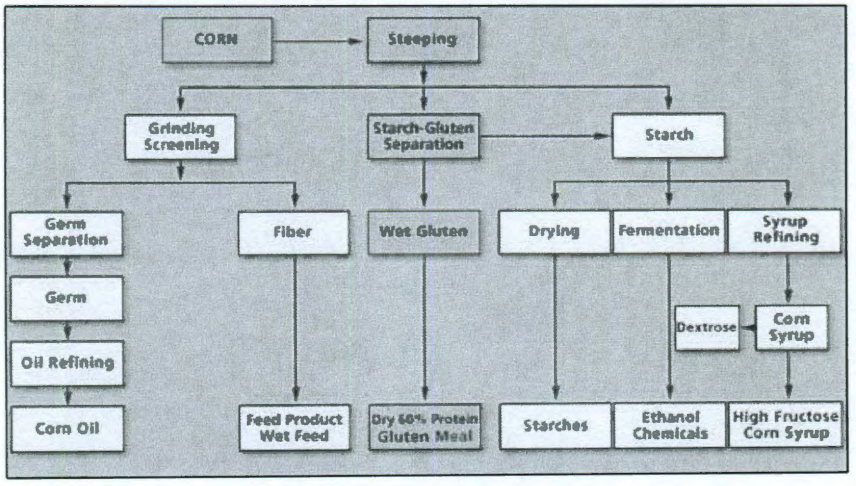



**Figure 7. Corn wet mill bio-refinery processes**

(source: Center for Transportation Analysis, Oak Ridge National Laboratory. http://cta.ornl.gov/bedb/biofuels/ethanol/The_Ethanol_Production_Process-Wet_Milling-Figure.xls)

## 1.3. FEASIBILITY OF REPLACING EXISTING PLATFORM CHEMICALS WITH BIO-BASED PLATFORM CHEMICALS

Another challenge for the production of bio-based chemicals is their integration into the existing chemicals market. Current platform chemicals such as ethylene, due to being derived from hydrocarbon feedstocks, have a low oxygen-to-carbon ratio - that is,

the carbon is highly reduced. However, many proposed bio-based platform chemicals, such as those in Figure 6, are derived from the carbohydrate portion of biomass, resulting in high oxygen-to-carbon ratios. Direct replacement of currently-used platform chemicals will require the removal of oxygen, an energy-dependent process that will add to the cost of producing bio-based chemicals, which is already struggling to be economically viable.

An alternative to direct replacement of platform chemicals is the functional replacement of existing petrochemicals. In order for this replacement to occur, the bio-based chemical must have functional properties similar to or better than the chemical it is replacing. Some proposed functional replacements with bio-based chemicals include ethyl lactate for replacing hydrocarbon solvents, 2,5-furandicarboxylic acid for terephthalic acid (monomer in polyesters), and glucaric acid for adipic acid (monomer in nylons) [13, 14]. This avenue of approach could yield additional benefits through improvements over existing chemicals, increased range of applications for certain chemical products, and the discovery of novel chemicals. However, in reality, finding new areas of chemicals applications has become increasingly difficult [6]. Consequently, functional replacements will most likely compete with existing chemical markets instead of creating a new market, as is the case with polylactic acid, a new bio-based polymer which has the features for replacing existing polymers but not enough to create a new market. In addition, there are high barriers to entry regarding the launch of a new industrial chemical, estimated to be US$1 billion covering both process and market development [15].

The best approach for a successful implementation of an industrial chemical industry based on biomass would most likely target both direct and functional

replacement of existing industrial chemicals. The bio-based platform chemicals listed in Figure 6, even with the aforementioned disadvantages, could still be involved in this type of approach, assuming highly efficient processes can be developed for them. However, unlike ethylene, propylene, and benzene, the 12 bio-based platform chemicals are much less reduced and do not have an extensive array of derivatization processes, restricting their impact to a limited number of products.

## 1.4. ADVANTAGES OF THE POLYKETIDE BIOSYNTHESIS PATHWAY

A more promising approach to bio-based chemicals is to move away from those that are carbohydrate-based to those that are polyketide/fatty acid-based. Chemicals resulting from this portion of biomass are composed of carbon that is in a more reduced form (lower oxygen-to-carbon ratio), making them better candidates for direct replacement since little to no energy will need to expended for their reduction. Secondly, the chemicals from polyketide/fatty acid synthesis (FAS) are produced from a single, cyclical metabolic pathway with the flexibility to generate various types and lengths of molecules that can carry an assortment of chemical functionalities in the form of controlled R-groups for the production of novel monomers. Lastly, fatty acid biosynthesis and its regulation have been studied extensively in *Escherichia coli*, providing key information needed for the rational design and engineering of the relevant metabolic pathways [16-18]. Thus, the polyketide/fatty acid biosynthesis provides an ideal framework for the production of a vast number of bio-based platform chemicals and subsequent manufacture of industrial chemical products. However, it is important to note that the biological end products derived from the FAS pathway, though they are highly reduced in nature, cannot directly replace existing platform chemicals as-is due to their remaining

carboxyl/carbonyl groups. To realize the production of bio-based industrial chemicals within the framework, they can act as close intermediate chemicals that can be readily converted through chemical catalysis into the desired platform chemical that can be integrated seamlessly into existing chemical industry infrastructure. Thus, a multi-disciplinary, collaborative research effort is required. Currently, this is spearheaded through the NSF Engineering Research Center for Biorenewable Chemicals (CBiRC), which has assembled from among eight U.S. and four foreign institutions a group of over 30 faculty knowledgeable in fields ranging from microbiology to chemical catalysis. Central to the overall research is the aforementioned idea of the optimized coupling of biocatalysis (use of an organism's metabolism for the production of a desired product) and chemical catalysis to connect the intermediate to the final platform chemical, which is shown in Figure 8. This structure is analogous to that of the petrochemical industry, where biocatalysis takes the place of feedstock conversion (crude oil refining), and chemical catalysis and polymerization constitute the derivatization processes.

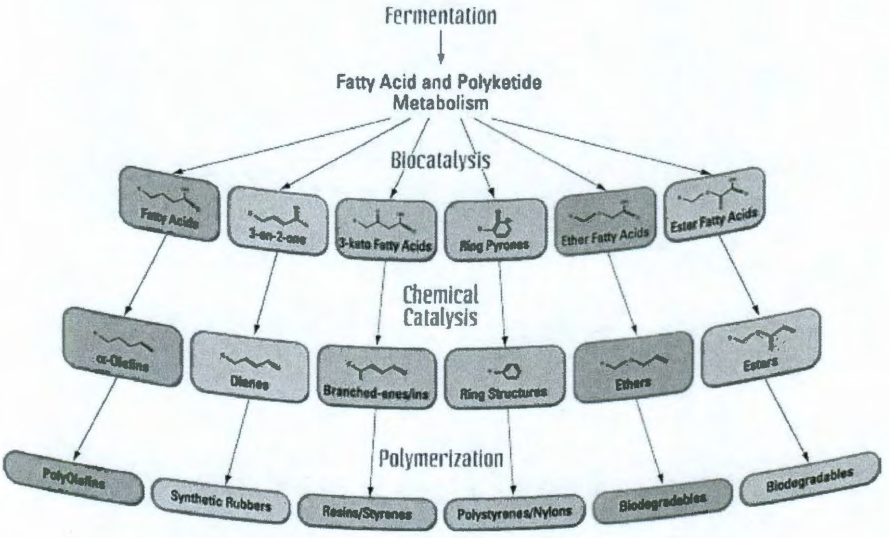



**Figure 8. CBiRC's engineering platform and various testbeds**

(source: NSF Engineering Research Center for Biorenewable Chemicals (CBiRC), http://www.cbirc.iastate.edu/researchdetails.asp)

The first row of compounds in Figure 8 are those that are known to be produced biologically in various organisms. The second row shows the corresponding final platform chemicals that result from modifications through chemical catalysis of the aforementioned compounds/intermediates. The third row shows the commodity chemicals that can be produced from the platform chemicals of the second row. In the proposed research, the focus will be on optimizing the first step of the process - the bioproduction of the intermediate chemicals in the first row.

## 1.5. THE GENERAL APPROACH AND TARGETS FOR PRODUCTION

In the area of biocatalysis, the first step is to identify and/or develop enzymes that will connect intermediates from the FAS pathway to the intermediate chemicals. This can be achieved through two strategies: (1) the structure-function relationship knowledge base can be used to develop new functionalities in existing biocatalysts, and (2) genomics

and metabolomics data from a wide variety of organisms can be used to identify novel enzymes that have been shown to produce these compounds from the FAS pathway. Once these enzymes are obtained, they will be introduced into *Escherichia coli*, a host organism that is ideal due to extensive knowledge of its metabolism, genetics, behavior, and the availability of a wide array of tools for its genetic modification. The next step is the development of design rules to systematically construct and optimize these strains for converting renewable feedstocks (such as glucose) into the desired products (second row, Figure 8). A systems approach will help guide the design of efficient microbial strains.

A primary long-term goal currently is the synthesis of compounds with 3-ene-2-one functionality, as shown in Figure 9.

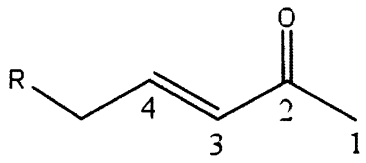



**Figure 9. The long-term target for the biological precursor to the ideal platform chemical**

This type of compound opens the door to a variety of possible platform chemicals through derivatization processes (chemical catalysis). For example: the carbon-to-carbon double bond can be selectively hydrogenated to form ketones; the carbonyl group can be hydrogenated to form unsaturated alcohols; the unsaturated alcohol can be dehydrated to form dienes. Dienes are an especially interesting product, as they can be used directly in the polymer industry and can be readily converted to olefins and diols, providing a great platform for a wide variety of chemical applications.

The short-term goal is a small step back from molecules with 3-ene-2-one functionality, and instead focuses on molecules with 2-one functionality (methyl

ketones). Methyl ketones can be used as platform chemicals themselves, and can also be readily converted to dienes which could directly replace existing platform chemicals such as ethylene and propylene.[19] Recently, two classes of methyl ketone synthases (MKS1 and MKS2) were isolated and characterized from wild and cultivated tomato as the enzymes responsible for generating methyl ketones from the FAS pathway [20, 21]. The second class of methyl ketone synthases, MKS2's, are thought to catalyze the first step of methyl ketone synthesis, which is the hydrolysis of the thioester bond of the 3-ketoacyl ACP intermediate in the FAS pathway, resulting in a 3-ketoacid. The first class, MKS1's, are thought to catalyze the next step, which is the decarboxylation of the 3-ketoacid into a methyl ketone.

Though it is not approached in the experimental portion of this study, another long-term goal is the bioproduction of ring pyrones. These can be converted through chemical catalysis into platform chemicals with ring structures, which have a wide variety of applications such as in the production of polystyrenes and nylons.[19] One pyrone of particular interest is triacetic acid lactone, or TAL (Figure 10). TAL can serve as a platform chemical by being readily converted through chemical synthesis to: (1) sorbic acid, a food and feed preservative, (2) 1,3-pentadiene, for resins, plastics, and adhesives, and (3) 3-pentene-2-one, a useful chemical in a variety of chemical syntheses. Though it is currently made industrially in five chemical steps starting with the pyrolysis of acetic acid [22, 23], there have been discovered three different biosynthetic routes to it from fatty acid synthesis intermediates.[24]

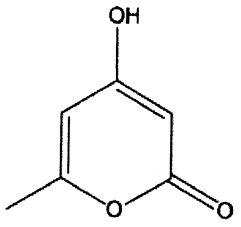



**Figure 10. Triacetic acid lactone**

While the discovery of new pathways for the synthesis of these target molecules is a critical first step, significant effort needs to be spent on incorporating and developing efficient microbial strains that will make producing these target molecules and the whole process economically viable. The specific objectives include: integration of new pathways into stable *E. coli* chassis, efficient pathway design to allow proper balance between cell growth and product formation, balanced carbon and co-factor flow, maintenance of robust performance even at high product titers, and eventual scale-up with industrial input.

## 1.6. HYPOTHESES

We hypothesized the successful production of methyl ketones through expression of heterologous proteins known to produce methyl ketones from fatty acid synthesis intermediates within wild tomato (*Solanum habrochaites*). Furthermore, we hypothesized increased titers and yields from growth conditions leaning microaerobic conditions, as this will increase the availability of reducing equivalents for use in fatty acid synthesis, the diversion of acetyl-CoA towards fatty acid synthesis through inactivation of the TCA cycle, while at the same time providing a minimal amount of oxygen to improve growth characteristics and to act as the external electron acceptor to complete the redox-balance

from glucose to methyl ketone. Finally, we hypothesized further increased titers and yields from inactivation of the mixed acid fermentative pathways that become active in microaerobic conditions, which not only compete for carbon with the fatty acid synthesis cycle but consume reducing equivalents required for the fatty acid synthesis reactions.

## 1.7. OBJECTIVES

The aim of this study is to enable the ability of *E. coli* to produce methyl ketones from glucose and apply metabolic engineering to increase the production efficiency. Prior studies on the production of platform chemicals have involved the over-production of relatively oxidized compounds from the carbohydrate portion of biomass, but no studies exist in the current literature concerned with engineering *E. coli* for the production of platform chemicals such as methyl ketones derived from intermediates in the polyketide portion of biomass. This study focuses on how heterologous proteins can be used in *E. coli* to synthesize platform chemicals such as methyl ketones and how its production can be optimized through balancing cell growth, carbon flow, and co-factor availability by consideration of aerobicity during growth and strategic pathway knockouts.

The specific objectives are described below and will be discussed in the body of this work:

- *Design of pathways leading to methyl ketones*: Initially, a systematic review of existing knowledge will be done regarding the biochemical reactions involved with the conversion of glucose to methyl ketones. This will lead to the identification of enzymes involved in the synthesis of methyl ketones (to be later introduced in various host strains), calculation of the overall balance from glucose to methyl ketones, and preliminary strategies for the optimization of the

recombinant pathways. Balance calculations will guide co-factor optimizations of producing strains. All of the above will provide the genetic design rules for metabolic engineering in order to achieve the overall objectives.

- *Identification of products*: SDS-PAGE analysis of whole cell extract will be used to confirm expression of methyl ketone synthases on the protein level, and methyl ketone production will be confirmed by GC-MS on the product level. Major and minor products will be identified to guide quantification of total methyl ketone production through GC-FID.
- *Strain characterization and establishment of optimal operating conditions:* Production strains will be assessed under a variety of operating conditions. External factors such as type of growth media and aerobicity will be varied. Internally, expression levels of methyl ketone synthases will be varied to find the proper balance between cell growth and product formation. Performance will be evaluated using fermentation profiling using HPLC and measurement of product formation through GC-FID. Metrics to be optimized are primarily titer and yield.
- *Strain improvement through genetic modification* : Following establishment of optimal operating conditions, potential metabolic engineering targets will be identified and executed. Effect of modifications will be evaluated using fermentation profiling using HPLC and measurement of product formation through GC-FID.

The proposed work will result in strains optimized for the production of the first polyketide/fatty acid-based platform chemicals, as well as increased understanding of the

pathways responsible for generating the intermediates to platform chemicals. In addition, this knowledge will greatly aid future research endeavors targeting other polyketide/fatty acid-based chemicals, bringing society closer and closer to a sustainable future.

# 3. LITERATURE REVIEW

The goal of this literature review is to cull: (i) all existing knowledge regarding the techniques for the metabolic engineering of *E. coli*, paying particular attention to the methodologies employed in the past for the production of novel compounds, and (ii) all knowledge on the protein, genetic, and biochemical levels pertaining to the synthesis of methyl ketones. Review of these two areas will lead to an initial set of design rules as well as a logical approach through which to execute on those.

## 3.1. METABOLIC ENGINEERING OF ESCHERICHIA COLI

Metabolic networks in cells are comprised of thousands of regulated and coupled chemical reactions that are driven by enzymes and membrane transport systems. Unfortunately, these networks as they have become evolved in nature are rarely optimized for practical applications. However, rapid improvements in molecular biology have given rise to technology that enables the genetic modification of cells to produce strains that have desirable properties.

Metabolic engineering is the application of molecular biology for the optimization of a specific network of biochemical reactions within an organism towards a specific goal. This goal is either the formation of desired products such as antibiotics, polymers, industrial chemicals, or the modification of cellular properties such as through the extension of metabolic capabilities. Inherent to these objectives, in the case of production of a metabolite, is the optimization of yield, productivity, and titers. In order to achieve these goals, several approaches (in terms of genetic modifications) have been used. For the increased production of an existing metabolite, typically the amplification of a gene or set of genes is performed to increase the flux along the biochemical reactions between

the substrate and the final product. Another method to enhance product formation is the removal or inhibition of an enzyme to knock out a competing pathway or one that leads to a toxic byproduct. For the production of novel compounds that are not normally produced by the cell, heterologous enzymes can be expressed to build the non-native pathways that result in the desired product. This strategy can also be used to extend substrate range or confer the ability to degrade inhibitory products, such as in the case of bioremediation. A combination of all these approaches may be needed to achieve the goal.

In order to apply these approaches in a rational and systematic manner, it is important to understand the properties of the pathway of interest from substrate to product, specifically the metabolic and information fluxes as well as their control. The use of analytical methods to reveal the underlying mechanisms of metabolic networks in combination with molecular biology techniques is what differentiates metabolic engineering from previous strain improvement efforts. Several techniques are used for this: (1) rigorous physiological studies, (2) metabolic flux analysis (MFA), (3) metabolic control analysis (MCA), (4) thermodynamics of the pathways, and (5) kinetic modeling. These techniques are used in metabolic engineering in order to determine the targets for modification as well as to correctly assess the consequences of a previous round of genetic modifications. In this study, due to the unknown regulatory patterns of fatty acid synthesis (incomplete knowledge), the lack of literature concerning MCA and MFA of the fatty acid synthesis pathway, these techniques were not used for rational modification in this study. In addition, the metabolic engineering was conducted was primarily through pathway deletions and the organization of the pathways were such that it did not

necessitate the need for MCA or MFA. Thus, the primary rationale technique for strain improvement was through analysis of the physiological studies.

### 3.1.1. Physiological studies

The combination of genetic work and classic cultivation studies are important for identifying unknown regulatory patterns and regulatory compounds. In addition, physiological studies allow the observation of the effects of genetic manipulations on growth and un-related systems. Normally, negative effects on growth can be attributed to increased metabolic burden on the cell by the modifications, but negative effects on un-related pathways are harder to anticipate. For example, one study amplified phosphenolpyruvate-forming enzymes and observed inhibition of heat-shock enzymes and nitrogen regulation [25]. Another study involved the use of continuous cultures to ascertain the regulatory aspect of an oft-used promoter (TAKA) of *Aspergillus oryzae*, showing both glucose repression and maltose induction [26, 27].

Prior engineering strategies for over-production of fatty acids have also been exclusively assessed through physiological studies in a relatively qualitative manner. Acetyl-CoA carboxylase, popularly over-expressed in studies where overproduction of fatty acids is the objective, was first over-expressed based on the belief that it was postulated to be the limiting step in fatty acid synthesis.[17] Only afterwards, after the effect of over-expression was confirmed to increase fatty acid production, was it established to be true. This type of rationale is in stark contrast to rationale applied to finding the appropriate targets for increasing glycolytic flux, which has been done in a more quantitative manner through MFA and MCA due to the knowledge of the kinetics of each enzyme along the pathway and large quantities of data from isotope labeling studies.[28]

The lack of data for fatty acid biosynthesis may be due to historically low interest in overproducing fatty acids. However, there has been largely increased interest in this topic in recent years,[6, 16, 17, 29-32] which may lead to more rigorous characterization of the fatty acid pathway in *E. coli* in the near future and a more quantitative basis upon which to rationalize further modifications.

### 3.1.2. Biosynthesis of industrial chemicals

Long before the earliest written records, mankind has employed micro-organisms for the production of useful compounds from raw materials; for example, in the making of beverages such as beer and wine, which employs yeast to convert sugars (from grains in beer, from grapes in wine) to alcohols, and in the making of bread, which employs yeast for leavening. However, actual knowledge of the metabolic processes involved in the production of alcohols and organic acids is fairly recent, with reports in the literature first appearing in the early 19$^{th}$ century. Lactic acid is speculated to be the first chemical to be purposefully produced on an industrial scale through fermentation [33]. In 1916, Chaim Weizmann patented the famous acetone-butanol-ethanol (ABE) fermentation process, which uses the fermentative metabolism of bacteria from the Clostridium Family to produce acetone, *n*-butanol, and ethanol from starch. This process played a vital role during World Wars I and II, as acetone was a key raw material in the manufacture of explosives. However, during the 1950s, the chemical synthesis of acetone from petroleum became a much cheaper alternative due to the sharp decline in the price of petroleum, and as of today there are no industrial plants running the ABE process.

On the other hand, scientific breakthroughs were being made in the field of molecular biology. The landmark discovery of the structure of DNA in 1953 by Watson

and Crick, followed by the development of recombinant DNA techniques by Cohen and Boyer in 1973, opened the door to intentional and purposeful genetic modifications, also known as genetic engineering [34, 35]. With this new tool, it was now possible to improve strains for the production of desired compounds and even introduce foreign DNA into them for the production of novel compounds.

Since the advent of genetic engineering, there has been increasing interest in engineering organisms for a multitude of applications, a significant one being the production of industrial chemicals. Petroleum-based conversion processes for the production of commodity organic chemicals are still highly efficient, being able to produce vast quantities of chemicals for relatively low cost due to economies of scale, flexibility, well-established nature, and current low price of oil. This economic efficiency creates a significant barrier to new production routes such as bioconversion. However, bio-production of organic chemicals comes with several advantages over petroleum-based production. Specifically, the ability to use biomass (plentiful lignocelluloses, agricultural products, wastes) allows for a sustainable way to produce chemicals, as opposed to using petroleum, which is established as a non-renewable resource. In addition, the environmental effects of bioconversion may be relatively mild. The use of biomass as a feedstock would reduce the net carbon dioxide production (as biomass consumes carbon dioxide), production of greenhouse gases during the bioconversion could be reduced if addressed correctly, wastes could be used as feedstock, and byproducts of bioconversion are typically benign. Unfavorable economics, the primary barrier for establishing a bio-based chemicals industry, can be overcome through the increased conversion efficiency attained through genetic/metabolic engineering as well as

the expected increases in the oil-to-biomass cost ratio. In a recent report on biomass technologies, the U.S. Department of Energy envisioned that by 2030, at least 25% of the raw materials for chemicals will come from renewable sources such as biomass [36].

There are three general bioconversion routes to organic chemicals: (1) the microbial fermentation of carbohydrates, (2) one-step enzymatic or microbial conversion of intermediates, and (3) enzymatic or microbial conversion followed by conventional catalytic conversion. The resultant biochemicals can enter the existing chemicals market through direct or indirect substitution. Direct substitution involves the production of an existing commodity chemical through a different, bio-based route; for example, as in the production of ethylene from ethanol. The advantage of direct substitution is that there is no need to develop a new market for the chemical since it already has established its utility. However, since new production routes for directly substituted chemicals competes with any pre-existent market share for that particular chemical, it must have an economic advantage over existing production routes. A review of the bioproduction of organic chemicals listed several candidates for direct substitution, which include acetaldehyde, acetic acid, acetone, acrylic acid, adipic acid, 1,3-butadiene, n-butanol, ethanol, ethyl acetate, ethylene, glycerol, isopropanol, maleic anhydride, methanol, methyl ethyl ketone, phenol, and propylene.[37]

In indirect substitution, a new chemical is produced that has a similar application to an existing chemical. In some cases, the new chemical may be better than the currently used chemical, in which case the current chemical is displaced. An advantage of indirect substitution is that it does not involve direct competition with existing chemicals, and thus bio-products have a higher chance at entering the market through this route. The

inherent disadvantage with this route is that new markets need to be established for the new chemical, and extensive product-development research may be required. An example of bioproducts that have entered through this route successfully are lactic acid copolymers such as polyhydroxybutyrate. The polymer sector of the organic chemicals industry is by far the largest in terms of annual production, and thus represents a significant market opportunity for potential bioproducts.[19] One example of an industrially produced bioproduct in this sector is polylactic acid polymer, produced by Cargill Dow LLC. PLA is considered to be a good alternative to petroleum-based plastics as there are several superior qualities to it such as biodegradability, biocompatibility, compostability, and low toxicity to humans over PET and other nylon-based plastics [38]. The first step is the microbial fermentation of bacteria to produce lactic acid from biomass sugars, which is then converted to a ring (lactide) and polymerized via chemical synthesis. Cargill Dow currently relies on corn grain as a source of glucose for PLA production, but research is on-going to develop fermentative organisms that allow a wider substrate range (e.g. the ability to convert pentose sugars) as well as a direct method to produce polylactide within the organism. Lactic acid itself is an important product of industrial importance, and thus various approaches have been attempted to improve its production within various organisms. Lactic acid bacteria, yeast, and fungal systems have been engineered to enhance lactic acid production. *E. coli* has also been metabolically engineered to enhance production of lactic acid from glucose, with most attempts involving the use of heterologous expression of lactic acid dehydrogenases in various expression systems coupled with the deletion of competing pathways, but in general the titers granted by these approaches are not competitive with lactic acid bacteria.[39]

Succinnic acid is another bio-product that has been of great interest in the past five to ten years. The value of succinic acid lies in the fact that many high-value derivatives can be easily attained through chemical synthesis.[19] Succinic acid is a major fermentative product of several bacteria including the obligate anaerobe *Anaerobiospirillum succiniciproducens* and facultative anaerobes belonging to the family *Pasteurellaceae*, such as *Actinobacillus succinogenes* and *Mannheimia succiniciproducens*.[40] *E. coli* is also known to produce succinic acid, but only as a minor fermentation product, with wild-type *E. coli* preferring the other fermentative products such as acetic acid, formic acid, lactic acid, and ethanol to retain its redox balance, necessitating the need to redirect metabolic fluxes for increased succinic acid production as well as reducing the formation of other metabolites. To this end, *ldhA* and *pfl* double mutants have been engineered to block formation of lactic, acetic, and formic acids.[41] Unfortunately, cell growth was relatively slow, the authors concluding that this was due to the inactivation of enzymes related to pyruvate dissimilation. Further enhancement has been attained by amplification of carbon-dioxide fixing anaplerotic pathways, catalyzed by phosphoenolpyruvate carboxylase[42, 43], malic enzyme, and pyruvate carboxylase[44].

## 3.2. THE BIOSYNTHESIS OF METHYL KETONES

Since the goal of the work is to engineer *E. coli* for the production of methyl ketones, the first step is to explore the current knowledge of the pathways and enzymes responsible for their production in order to identify potential enzymes for introduction into the *E. coli* chassis. The pathway information will provide the way to connect a native metabolite in *E. coli* to methyl ketones, and any related enzyme sequence information

will provide the starting point in the genetic engineering required to introduce those necessary biochemical reactions in *E. coli*.

Aliphatic methyl ketones are a class of organic chemicals whose natural origin has been studied and firmly established over many years. The occurrence of methyl ketones in nature was first reported in 1858 when Williams discovered 2-undecanone as the primary constituent of oil from *Ruta graveolens,* a plant species [45]. Later, 2-undecanone was also found to be present in other essential oils of plants [46]. Afterward, many reports confirmed plants as producers of methyl ketones, with 2-nonanone, 2-undecanone, and 2-tridecanone found in various quantities in a great many plant and seed oils [47-49]. In some cases, their associated secondary alcohols (2-nonanol, 2-undecanol, 2-tridecanol) occur together, and are thought to be the reduced equivalents of their corresponding methyl ketones [50].

However, plants are not the only producers of methyl ketones. They were also found to be in cheeses, milk, and butter by food scientists whose primary interests were to identify the source of volatile flavors [51]. In fact, methyl ketones are ubiquitous compounds known to be produced by almost all walks of life, with at least one example in five of the six kingdoms of life (Animalia, Plantae, Fungi, Archaea, and Bacteria). This section aims to present current knowledge of the pathways, enzymes, and other relevant information pertaining to the production of methyl ketones in nature.

### 3.2.1. Overview of the main pathways to methyl ketones

There are two main pathways through which methyl ketones are synthesized. One pathway involves an intermediate from the fatty acid synthesis cycle (Figure 11), 3-ketoacyl-ACP, which undergoes hydrolysis of the thioester bond catalyzed by a

thioesterase to form 3-ketoacid, which is subsequently irreversibly decarboxylated by a beta-decarboxylase to produce an odd-numbered methyl ketone (the intermediates of the fatty acid synthesis cycle are even-numbered). This is what is known to occur in plants [20]. The other pathway, which occurs mainly in bacteria and fungi, involves 3-ketoacyl-CoA, an intermediate from fatty acid degradation (Figure 12), producing methyl ketone via a similar thioesterase and beta-decarboxylase.

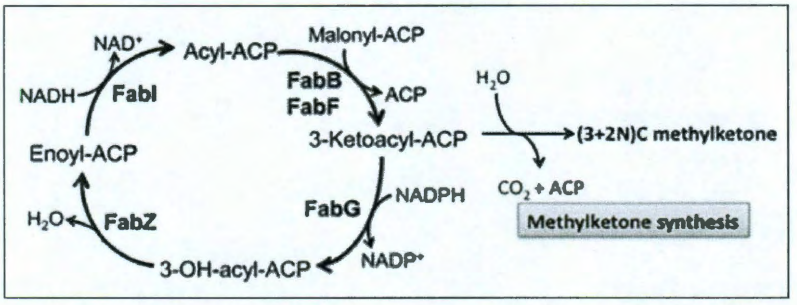



**Figure 11. Methyl ketone synthesis from the fatty acid biosynthesis pathway**

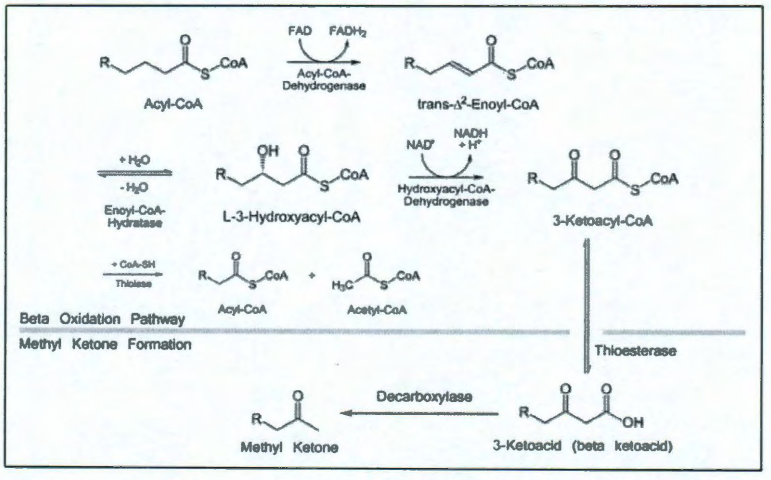



**Figure 12. Methyl ketone synthesis from the fatty acid degradation pathway**

As is evident from the figures and pathway description, the methyl ketone precursors of both pathways differ only in their acyl group carrier. Another similarity is that both involve cleavage of the thioester bond and beta-decarboxylation.

Indeed, there are at least three other ways bacteria produce methyl ketones not involved in either the production or degradation of fatty acids. Similar to the pathways mentioned earlier, one way is by decarboxylation of beta-ketoacids that are not the result of thioesterases. Many *Clostridia* species and some *Pseudomonas* produce ketones through beta-decarboxylation, most notably acetone from acetoacetate [52-54]. Another way bacteria produce methyl ketones is by oxidation of secondary alcohols [55]. Alkane degradation pathways in some bacteria also produce methyl ketones as intermediates in the pathway by converting alkanes into secondary alcohols via monooxygenase and subsequent oxidation by secondary alcohol dehydrogenase. Though five pathways to methyl ketones have been mentioned, there are only three types of immediate precursors to methyl ketones, which are shown in Figure 13. Specific details regarding these pathways and responsible enzymes will be found in the following sections.

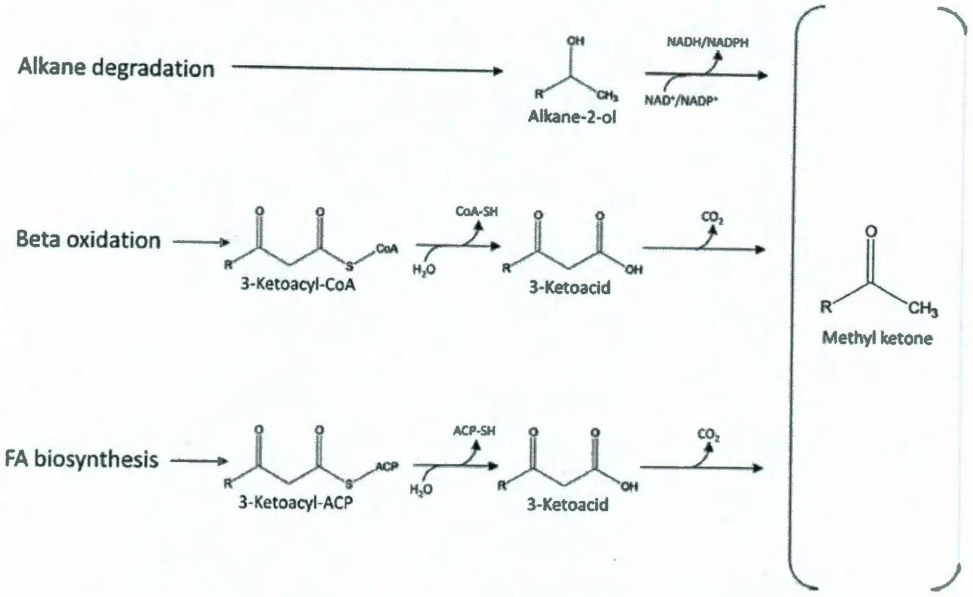



**Figure 13. Main precursors and reactions for producing methyl ketones.**

## 3.2.2. Methyl ketones in Plantae

Many plants produce volatile essential oils synthesized and secreted by special glandular hairs on their leaves and stems called "trichomes". Many of the constituents of these oils are used by the plants to help ward off detrimental pests and contain methyl ketones of various lengths. In particular, medium-length methyl ketones such as 2-undecanone are very effective at warding off insects and are found to be produced by many species in the plant kingdom [56, 57]. Among these, one of the most prominent sources of methyl ketones are the glandular trichomes of the wild tomato species *Solanum habrochaites* [57]. As this species is by far the most well-studied regarding synthesis of methyl ketones (in fact, no other study exists on the subject of the identification of key enzymes in methyl ketone synthesis), this section will primarily focus on knowledge so far regarding this species.

In 2005, an enzyme expressed in *S. habrochaites* (accession PI126449) leading to the synthesis of methyl ketones was isolated and characterized for the first time [20]. This was done through a comparison of the gland EST databases between an accession of *S. habrochaites* (PI126449) that was a known producer of methyl ketones, and an accession (LA1777) that does not produce methyl ketones. An EST is an "Expressed Sequence Tag", a short sub-sequence of a cDNA sequence used for the identification of gene transcripts. Thus, through this comparison, a cDNA was identified that was abundant in PI126449 but rare in LA1777, and the enzyme encoded by the cDNA was designated Methyl Ketone Synthase 1 (MKS1). Further analysis showed the sequence to be similar to plant thioesterases (Figure 14), and levels of the MKS1 transcript, protein, and enzymatic activity in a variety of accessions and plant organs correlated with the amount of methyl ketones present, implicating MKS1 as a key enzyme in methyl ketone synthesis. The MKS1 protein was then expressed in *E. coli* and purified from whole cell extract to catalyze *in vitro* reactions using $C_{12}$, $C_{14}$, and $C_{16}$ 3-ketoacyl-ACP substrates, which led to the production of $C_{11}$, $C_{13}$, and $C_{15}$ methyl ketones respectively, suggesting that MKS1 can act as both a thioesterase and a decarboxylase, though the turnover rate was low. Sequence similarity showed that MKS1 is a member of the large alpha/beta-hydrolase-fold superfamily of esterases [58]. In this particular study, *in vivo* methyl ketone production in their recombinant *E. coli* strain was not measured, possibly due to the lower-than-detectable amounts of methyl ketones produced by their recombinant strain, or the lack of a method to extract methyl ketones from *E. coli* at the time.

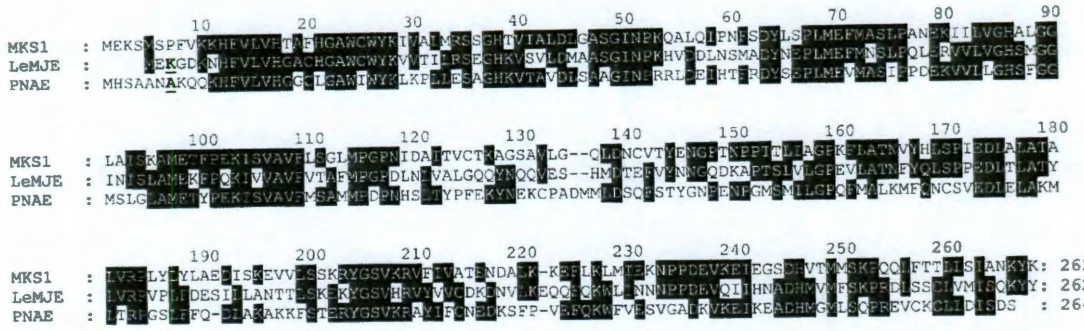


**Figure 14. Amino acid sequence alignment of MKS1 and two plant esterases [20]**

(LeMJE refers to methyl jasmonate esterase from *L. esculentum*, and PNAE referes to polyneuridine aldehyde esterase from *R. serpentina*).

In 2009, the same group published another study identifying two additional methyl ketone synthases, designated SHMKS2 and SLMKS2, from second-generation interspecific populations of *S. habrochaites* (wild tomato) and *S. lycopersicum* (cultivated tomato) respectively that were observed to have high methyl ketone content in their trichomes. Crossing the wild tomato species which can produce methyl ketones with the cultivated tomato species which cannot, allowed a segregation analysis that concluded a multiple quantitative trait loci requirement for the production of methyl ketones. In addition, five plants with high levels of 2-tridecanone production and five plants with no detectable levels of 2-tridecanone (all from the F2 generation) were chosen for a transcriptome analysis, where the RNA from the trichomes was extracted, reverse transcribed, labeled, then hybridized to a microarray containing tomato genes. Table 2 shows the significant differentially expressed genes between the two bulks. One gene that was especially over-expressed (337-fold more) in the methyl ketone producing F2 plants was designated *SHMKS2*, and a similar transcript derived from the cultivated tomato species was 7.5-fold more highly expressed in the non-producing plants and designated *SLMKS2*.

**Table 2. Transcriptome analysis of genes that were differentially expressed in producing- and non-producing- interspecific F2 strains [21].**

| Gene Code[a] | Annotation | Ratio[b] |
|---|---|---|
| DN167657 | A protein related to a *Pseudomonas* thioesterase (Sh allele) | +336.6 |
| AI779239 | rRNA-16S ribosomal RNA | +62.6 |
| AF230371 | Allene oxide synthase | +46.7 |
| BI925004 | Plasma membrane intrinsic protein | +39 |
| BI931228 | Unknown | +27.3 |
| AW616884 | Dehydrodolichyl diphosphate synthase | +24.1 |
| DN169296 | DNA repair protein RAD23 | +18.3 |
| DB719610 | Calcium-binding EF hand family protein | +15.5 |
| DN168712 | Rubisco small subunit 1A | +14.7 |
| DN169129 | Major latex protein-related (Sh allele) | +12.9 |
| AW039905 | Peroxisomal protein involved in the activation of fatty acids | +11.8 |
| AI777019 | Unknown | +11.2 |
| AW615872 | Glycosyltransferase family 14 protein | +11.1 |
| BF097749 | Mitochondrial 26S ribosomal RNA protein | +9.2 |
| BW688217 | Unknown | +9.1 |
| BM412813 | Methyltransferase family 2 protein | +8.8 |
| DN170232 | Protein kinase | +8.4 |
| DN171038 | Casein kinase 1 protein family | +7.2 |
| DB683900 | X-Pro dipeptidase | −7.1 |
| BG131749 | A protein related to a *Pseudomonas* thioesterase (Sl allele) | −7.5 |
| AI772024 | Unknown | −7.6 |
| DB722221 | Unknown | −7.7 |
| DN168641 | PSII oxygen-evolving complex 23 | −7.8 |
| ES893822 | Cell wall protein precursor | −7.9 |
| BG643000 | Phospholipase A2$\beta$ | −8.0 |
| BW690350 | GRAM domain-containing protein/abscisic acid-responsive protein-related | −8.3 |
| AW034502 | Cytochrome P450, putative | −8.9 |
| BG1237766 | Aldo/keto-reductase family protein | −9.5 |
| BI928231 | NAD-dependent epimerase/dehydratase family protein | −10.1 |
| BI932160 | UDP-glucoronosyl/UDP-glucosyl transferase family protein | −10.6 |
| BG128416 | Unknown | −10.9 |
| AW624755 | Major latex protein-related | −11.9 |
| ES896328 | Chaperonin | −12.5 |
| BE434841 | 3-Ketoacyl-ACP synthetase 2 nuclear gene | −15.2 |

[a]Corresponding GenBank accession numbers with the highest similarity are shown. [b]Ratio depicts the difference in average ratios of high-MK over low-MK bulks in four hybridizations for all of the probes that represent the same sequence.

The MKS2 protein is approximately 50-70% identical to various plant proteins whose functions are unproven, encoded by genes in the *Arabidopsis thaliana* and *Oryza sativa* genomes, and by ESTs found from the angiosperm family and white spruce *Picea glauca*. The sequence similarity has implied that the MKS2 protein is part of the 4-hydroxybenzoyl-CoA thioesterase subfamily of hot-dog fold enzymes [59]. The protein sequence alignment of SLMKS2 and SHMKS2 to related plant proteins is shown in Figure 15.

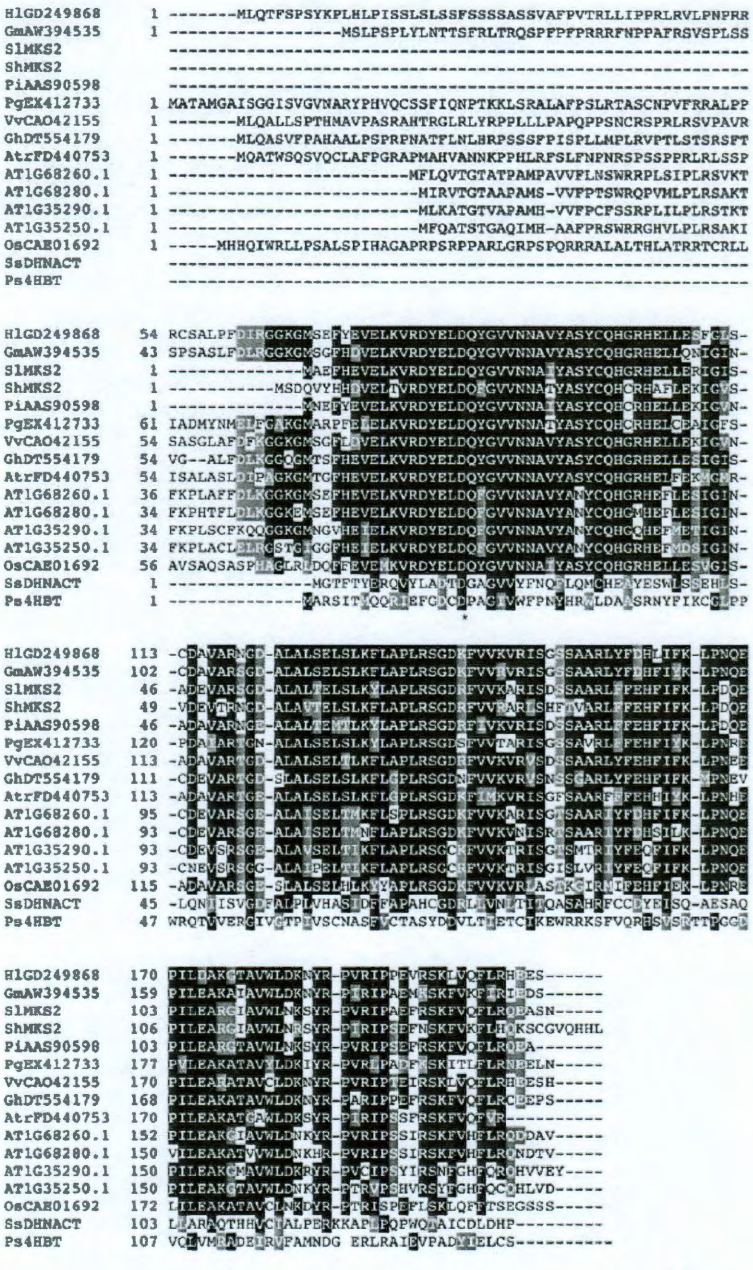


**Figure 15. Protein sequence alignment of SLMKS2 and SHMKS2 with related proteins [21].**

The white letters on black background indicate conserved amino acid substitutions. The asterisk refers to the catalytic Asp residue identified in 4-hydroxybenzoyl-CoA thioesterase. At, Arabidopsis thaliana; Atr, Amborella trichopoda; Gh, Gossypium hirsutum; Gm, Glycine max; Hl, Humulus lupulus (cv Phoenix); Os, Oryza sativa; Pa, Prunus armeniaca; Pi, Petunia integrifolia subsp. inflata; Pg, Picea glauca; Ps, Pseudomonas sp. (strain CBS-3); Sh, Solanum habrochaites; Sl, Solanum lycopersicum; SsDHNACT, Synechocystis sp. PCC6803 1,4-dihydroxy-2-naphthoyl-CoA thioesterase; Vv, Vitis vinifera. Accession numbers are as follows: Atr, FD440753; Gh, DT554179; Gm, AW394535; Hl, EX521228; Os, CAE01692; Pi, AAS90598; Pg, EX412733; SsDHNACT, NP442358; Vv, CAO42155.

These enzymes also had similarities to other thioesterases found in other Kingdoms. A BLAST algorithm was performed to search for proteins with similar sequences outside of plants, and it was found that MKS2 is 34% identical and 48% similar to more than 131 aligned residues of a putative thioesterase expressed by *Thermus thermophilus*. This putative thioesterase is also 26% identical and 51% similar to 4-hydroxybenzoyl-CoA thioesterase, supporting the notion of the homology of MKS2 to 4-hydroxybenzoyl-CoA thioesterase. The study implied the cooperation of both the MKS2-class of enzymes and the previously identified MKS1-class of enzyme with regards to methyl ketone production from 3-ketoacyl-ACP. Though unconfirmed, they suggest a mechanism in which, similar to that of bacterial methyl ketone synthesis, one enzyme (putatively, both SHMKS2/SLMKS2) hydrolyzes the thioester bond to produce 3-ketoacid, which gets decarboxylated by MKS1. Recent, unpublished data showing high in vitro decarboxylase activity of MKS1 and low decarboxylase activity of MKS2 seems to support this hypothesis.

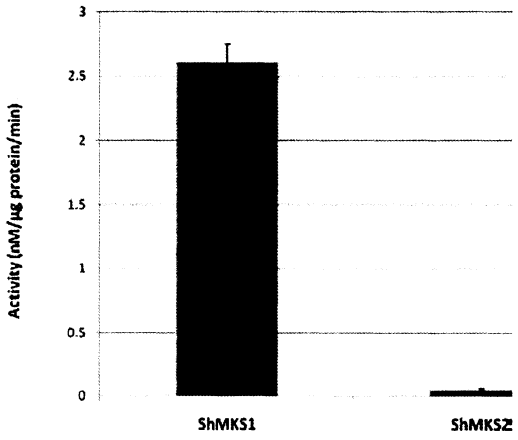



**Figure 16. In vitro assays for decarboxylase activity of SHMKS2 and MKS1 (Pichersky, unpublished)**

### 3.2.3. Methyl ketones in Animalia

Many mammals produce acetone during fatty acid degradation by decarboxylation of acetoacetic acid, the remaining intermediate in the last cycle of beta-oxidation. This precursor can also come from the degradation of the amino acids leucine, isoleucine, lysine, phenylalanine, and tyrosine. Thus far, acetone is the only methyl ketone that has been detected in animal tissue. Though the enzymes responsible for this conversion have been isolated, acetone does not fit within the particular framework of this study.

Methyl ketones are also found in insects, where they are found as components in volatile secretions such as pheromones. 2-Heptanone has been found in honeybees and in ants *Atta texana* [60, 61]. 2-Tridecanone, 2-pentadecanone, and 2-heptadecanone have been found in the secretions of the ants *Acanthomyops claviger* and in *Lasius fuliginosus* [62, 63]. In general, many other insect secretions have been found to contain methyl ketones spanning 2-butanone to 2-heptadecanone. However, these studies have only been limited to the identification of the secretion constituents, and no enzyme isolation or characterization studies have been done regarding the methyl ketones found in the secretions, and the reactions leading to the synthesis of these compounds have not been reported.

### 3.2.4. Methyl ketones in Fungi

The study of methyl ketone production by fungi precedes the analogous studies in bacteria, receiving much attention since the early 1900s. Based on isotope studies and experiments with enzyme preparations from various fungi in the 1960s, it is now widely

accepted that methyl ketones produced from fatty acids by molds proceeds via the fatty acid degradation pathway (Figure 17) [64]. The process is widespread among filamentous fungi; one study found 30 out of 38 fungi tested converted fatty acids to methyl ketones. Another study showed 9 out of 11 *Aspergillus* species and 9 out of 12 *Penicillium* species also carried out the transformation [65]. In general, most literature on the subject focuses on these two genera.

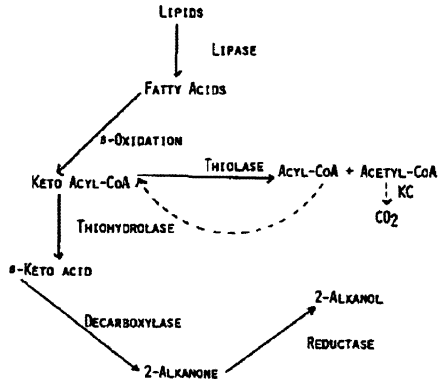



**Figure 17. Proposed pathway to methyl ketone from lipids or fatty acids in fungi. Broken line indicates possible inhibition of thiolase (thioesterase) by acyl-CoA [66]**

In these molds, beta-ketoacyl-CoA thioesterase activity (labeled thiohydrolase in Figure 17) greatly exceeds that of thiolase, and the beta-ketoacid decarboxylase apparently causes rapid decarboxylation, preventing the accumulation of beta-ketoacids. Though studies to date have shown beta-ketooctanoyl-CoA to be the preferred substrate for the thioesterase, methyl ketones of varying chain lengths are formed from long-chain fatty acids at successive cycles of fatty acid degradation [67]. Thus, of the methyl ketones found, heptan-2-one is usually the most abundant, followed by nonan-2-one, penta-2-one, and undecan-2-one (Figure 18). However, as aforementioned, the genes for the

thioesterase and decarboxylase have yet to be discovered, and all studies thus far have been limited to assays using enzyme isolates. A GenBank and Brenda search for thioesterases acting on ketoacyl-CoA and decarboxylases acting on beta-ketoacids yield no results, not only for *Penicillium* and *Aspergillus* but all the organisms in the database. Considering the long history of research on methyl ketone-producing molds, it seems there is little motivation in the area of food science to identify the genes responsible for coding the beta-ketoacyl-CoA thioesterases and beta-ketoacid decarboxylases.

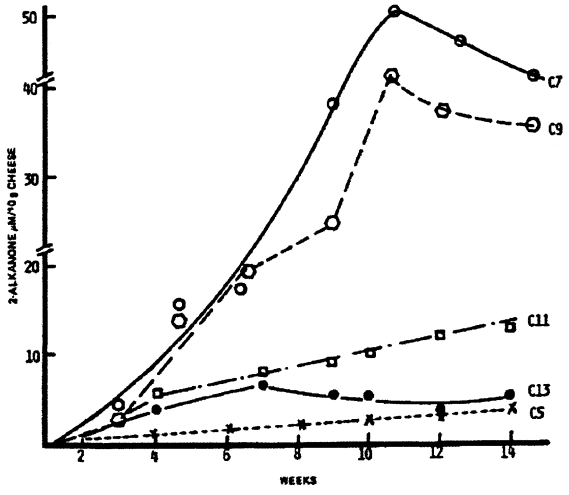



**Figure 18. Rate of accumulation of different methyl ketones during ripening of cheese samples [66].**

## 3.2.5. Methyl ketones in Archaea/Bacteria

This section will focus on current knowledge of enzymes responsible for methyl ketone production in bacteria. Many bacteria have been found to produce a variety of methyl ketones from acetone to 2-pentadecanone. However, there is a certain disparity between the amount of species that are known to produce methyl ketones, and the amount of species for which the responsible enzymes have been characterized (the latter amount

being much smaller). Thus, there will be more focus on the bacterial species for which the enzymes have been characterized and less on those which have currently only been shown to produce methyl ketones i.e. through cell-free extract activity assays.

The first pathway of Figure 13 shows methyl ketones derived from secondary alcohols such as 2-propanol, 2-butanol, 2-pentanol, and 2-hexanol by secondary alcohol dehydrogenase. Many species have been found to show this type of activity, and include: *Brevibacterium* sp., *Nocardia paraffinica*, *Pseudomonas fluorescens*, *Pseudomonas maltophilia, Pseudomonas multivorans, Mycobacterium sp., Candida utilis*, *Hansenula polymorpha*, *Pichia* sp., *Torulopsis* sp., *Kloeckera* sp., *Thermoanaerobium sp.*, and some methylotrophic bacteria including *Methylosinus trichosporium, Methylosinus sporium, Methylocystis parvus, Methylomonas methanica, Methylomonas albus, Methylobacter capsulatus, and Methylobacterium organophilum* [55, 68-71]. All the aforementioned species' secondary alcohol dehydrogenases are NAD-dependent with the exception of *Thermoanaerobium*'s, which is NADP-dependent, and almost all are specific to alcohols of six carbons or less, with the highest rates of oxidation for 2-propanol and 2-butanol. Most studies of these organisms have been limited to enzyme assays of cell-free extracts or purified enzyme, but a select few have been investigated further. One of the well-studied species among these is *Pseudomonas fluorescens*, with its genome being fully sequenced in 2005 [72]. It is a Gram-negative, obligate aerobe that primarily resides in soil and water, but can be found in many other places due to its versatile metabolism. In *P. fluorescens* and all the other preceding organisms, the secondary alcohol dehydrogenase specific to alkan-2-ols is important as it catalyzes a step in their alkane degradation

pathway, which converts alkanes to free fatty acids for beta oxidation, as shown in pathway 2 of Figure 19 [73].

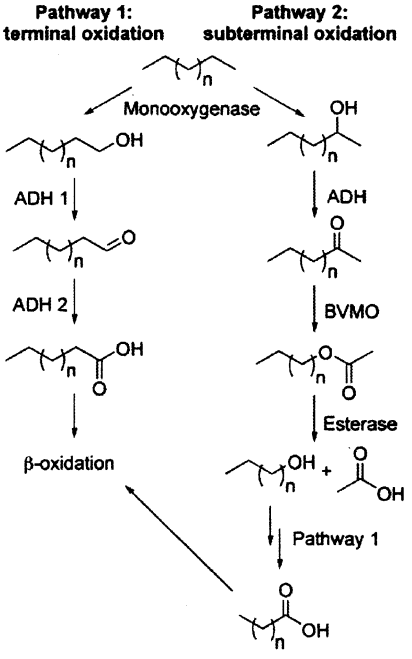



**Figure 19. Alkane degradation pathway [73]**

In a recent study, *Escherichia coli* was successfully engineered to degrade secondary alcohols 2-decanol and 2-dodecanol into 1-octanol and 1-decanol respectively[73]. The genes of the secondary alcohol dehydrogenase (*adhF1*), Baeyer–Villiger monooxygenase or BVMO (*bmoF1*), and esterase (*estF1*), the sequences of which can be found in NCBI's GenBank, were sub-cloned into a vector and recombinantly expressed in *E. coli*. The data pertaining to the conversion of secondary alcohols to primary alcohols in the engineered *E. coli* strain is summarized in Figure 20.

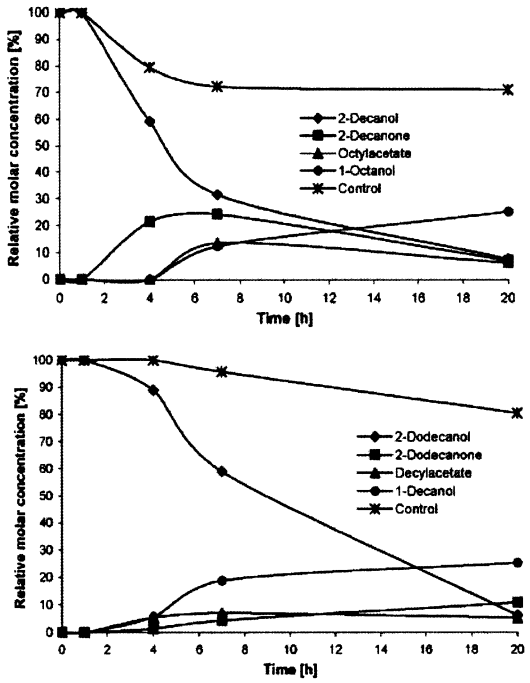



**Figure 20. Conversion of decan-2-ol and dodecan-2-ol in the engineered E. coli strain [73]**

Though the engineered *E. coli* produced primary alcohols as the end product instead of ketones, it can be assumed from the study's data that recombinantly expressing only the dehydrogenase would yield an *E. coli* that produces methyl ketones from secondary alcohols. With the addition of the monooxygenase that subterminally oxidizes alkanes, the *E. coli* may be able to produce methyl ketones from alkanes.

The second pathway of Figure 13 shows methyl ketones derived from decarboxylation of beta-ketoacids. The beta-ketoacids used in bacteria that produce methyl ketones this way are thought to be generated by the hydrolysis of the thioester bond in beta-ketoacyl-CoAs (an intermediate in the fatty acid degradation pathway) by thioesterase, and this pathway is shown best in Figure 12 [54]. The bacterial species that are known to include both these reactions are many. Cyanobacteria and actinomycetes

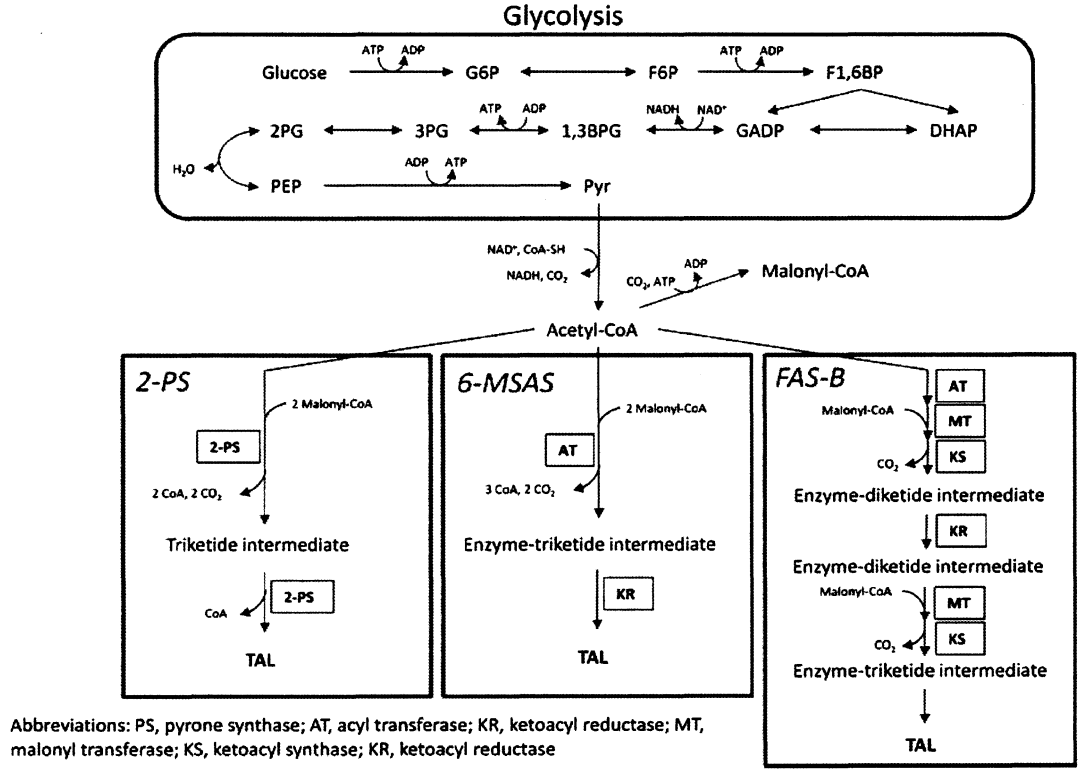
produce 3-pentanone and 3-octanone [74]. Certain *Serratia* strains produce 2-nonanone, 2-undecanone, and 2-dodecanone [75]. *Staphylococcus carnosus* produces 2-pentanone from ethylbutyryl acetate or acetone from acetoacetate [76]. Similarly, *Clostridium acetobutylicum* has a well-characterized acetoacetate decarboxylase (known gene sequence) that catalyzes the formation of acetone [52]. Unfortunately, many of these studies (with the exception of the *Clostridia* study) have been limited to enzyme assays using cell-free extracts. In general, the genes responsible for coding beta-ketoacid decarboxylases in both bacteria and fungi are unknown, as a quick search of GenBank's database will show.

The third pathway in Figure 13 is not yet known to occur in any bacteria, though it has been found to occur in plants [20].

## 3.3. TRIACETIC ACID LACTONE BIOSYNTHESIS

Though not considered in the experimental portion of this study, tracetic acid lactone is still worth mentioning due to it being a polyketide-based platform chemical and also because it can benefit from the metabolic engineering strategies and knowledge gained from the methyl ketone study. Triacetic acid lactone is currently manufactured industrially through a series of five steps starting with the pyrolysis of acetic acid.[23] Microbial synthesis of TAL has been found to occur in *Gerbera hybrida*, through an enzyme isolated as 2-PS encoded by the g-2-ps-1 gene.[24] In addition, the genetic modification of 6-MSAS, encoding the enzyme 6-methylsalycilic acid synthase from *Penicillium patulum*, and *fasB*, encoding the *Brevibacterium ammoniagenes* fatty acid synthase B, has been determined to have TAL-synthesizing activities that are not found in

their native forms. The pathways from each of these enzymes are shown in Figure 21. Synthesis of TAL in *E. coli* has only been preliminarily studied, through the heterologous expression of only 2-PS in a wild-type host. Unfortunately, the study concluded that the production of TAL was insignificant and thus did not report any additional information regarding TAL production in *E. coli* from the heterologous expression of these enzymes. However, applying metabolic engineering has the potential to enhance TAL production within *E. coli*, and thus is suitable as a future work outside of this study.




Abbreviations: PS, pyrone synthase; AT, acyl transferase; KR, ketoacyl reductase; MT, malonyl transferase; KS, ketoacyl synthase; KR, ketoacyl reductase

**Figure 21. TAL synthesis routes**

## 3.4. FATTY ACID BIOSYNTHESIS

The understanding of the fatty acid biosynthesis cycle is critically important in the context of producing methyl ketones as it is the direct supplier of the methyl ketone precursor, 3-ketoacyl-ACP. Fortunately, there is much literature on the subject, and in recent years, several groups have introduced a variety of modifications that have increased yields of fatty acids. The following section reviews some of these developments.

### 3.4.1. The reaction pathways

Figure 22 shows the two pathways regarding the initiation reactions to produce fatty acid chains, separated for clarity. The overall balance from glucose is shown in Figure 31. As in plants, fatty acids in *E. coli* are produced exclusively from acetyl-CoA with NADPH as the source of reducing equivalents. The fatty acid synthase system in *E. coli* is of the "type II", also known as dissociated fatty acid synthase systems. As opposed to the mammalian fatty acid synthesis which is catalyzed by individual domains of a very large polypeptide, the reactions in bacteria are catalyzed by separate enzymes encoded by unique genes. However, there are usually multiple proteins that catalyze the same chemical reaction, though the difference in substrate specificities allows the regulation of the range of products produced by the pathway.

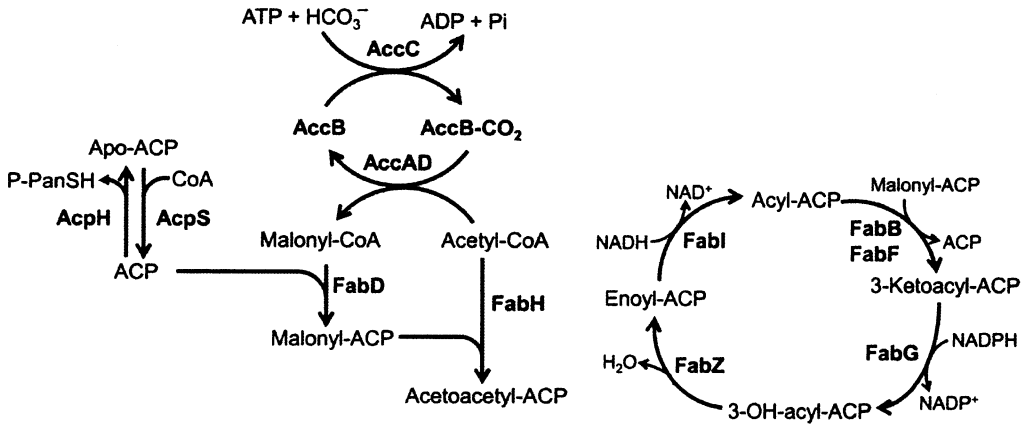



**Figure 22. Fatty acid biosynthesis separated into initiation and elongation cycles.**

(source: *ecosal.org*)

Acetyl-CoA carboxylase (ACC) catalyzes the first committed, irreversible step of the pathway, forming malonyl-CoA in an ATP-dependent manner [29]. The ATP-dependent carboxylation adds one more carbon, which is lost later during condensation with the growing acyl-ACP. It is a two-step reaction, involving the enzyme prosthetic group biotin. The ATP-dependent carboxylation of biotin occurs at one active site, followed by the transfer of the carboxyl to acetyl-CoA at a second active site. Biotin is connected to the enzyme by an amide bond between the carboxyl of the biotin side chain and the amino group of a lysine residue, which makes the biotin into an arm that can "swing" between the two active sites [77]. The overall reaction is as follows:

$$\mathbf{HCO_3^- + ATP + acetyl\text{-}CoA \Rightarrow ADP + P_i + malonyl\text{-}CoA}$$

The malonyl-CoA is then made available for its condensation with acetyl-CoA by FabD, a malonyl-CoA:ACP transacylase. This condensation reaction forms the initial molecule that enters the elongation reactions.

The elongation reactions are shown to the left of Figure 22. The first step involves condenstion of malonyl-ACP with the growing acyl-ACP chain by a beta-ketoacyl-ACP synthase. This is the only irreversible reaction in the elongation process, and thus it is key in regulating the product spectrum of the pathway [17]. The resulting 3-ketoacyl-ACP is reduced by an NADPH-dependent beta-ketoacyl-ACP reductase, and then dehydrated by beta-hydroxyacyl-ACP dehydratase. A final reduction takes place, catalyzed by enoyl-ACP reductase to form acyl-ACP, which can undergo another round of the elongation cycle. It is known that these reactions can be catalyzed by a number of isozymes differing in their substrate specificities, though at the moment only the initial condensation reaction has been found to be catalyzed by more than one enzyme. However, it is probable based on genetic and biochemical evidence that at least three beta-ketoacyl-ACP synthases and at least two beta-hydroxy-acyl-ACP dehydratases exist [17].

The final step in the fatty acid biosynthetic pathway for the production of fatty acids is the cleavage of the thioester bond of acyl-ACP by thioesterases. Again, there are at least two thioesterases expressed in *E. coli* that act upon acyl-ACPs, but differ in their substrate specificities. The genes known to express enzymes involved in fatty acid synthesis are listed in Table 3.

Table 3. Fatty acid synthesis-related proteins and their known functions [17].

| Gene | Protein or function affected | Map position | Mutant biochemical or growth phenotype |
|---|---|---|---|
| *aas* | Acyl-ACP synthetase | 61 | Accumulation of lysophosphatidyl ethanolamine |
| *accA* | Acetyl-CoA carboxylase | 4 | Carboxyltransferase α subunit |
| *accB* | Acetyl-CoA carboxylase | 71 | BCCP subunit |
| *accC* | Acetyl-CoA carboxylase | 71 | Biotin carboxylase subunit |
| *accD* | Acetyl-CoA carboxylase | 50 | Carboxyltransferase β subunit |
| *acpP* | ACP | 24 | ACP structural gene |
| *acpS* | ACP synthetase | 43 | Accumulation of apo-ACP |
| *cfa* | Cyclopropane fatty acid synthase | 36 | Cells lack cyclopropane fatty acids |
| *fabA* | β-Hydroxydecanoyl ACP dehydrase | 22 | Unsaturated fatty acid auxotroph |
| *fabAup* | | | Overproduction of saturated fatty acids |
| *fabB* | β-Ketoacyl-ACP synthase I | 50 | Unsaturated fatty acid auxotroph |
| *fabD* | Malonyl-CoA:ACP transacylase | 24 | Ts[a] auxotroph requiring both saturated and unsaturated fatty acids |
| *fabE* | Acetyl-CoA carboxylase | 71 | Now designated *accB* |
| *fabF* | β-Ketoacyl-ACP synthase II | 24 | Altered thermal regulation |
| *fabG* | β-Ketoacyl-ACP reductase | 24 | |
| *fabH* | β-Ketoacyl-ACP synthase III | 24 | |
| *fadR* | Transcriptional regulator of *fabA* | 26 | |
| *fatA* | | 69 | Utilization of *trans*-unsaturated fatty acids |
| *lpxA* | UDP-Glc-NAc acyltransferase | 4 | Accumulation of lipid A |
| *orf-17* | Putative dehydrase | 4 | Unknown |
| *plsB* | *sn*-Glycerol-3-phosphate acyltransferase | 92 | Glycerol-3-phosphate auxotroph |
| *plsC* | 1-Acylglycerol phosphate acyltransferase | 65 | |
| *plsX* | Unknown | 24 | Required for PlsB$^-$ phenotype |
| *tesA* | Thioesterase I | 12 | Defective enzyme activity |
| *tesB* | Thioesterase II | 10 | Defective enzyme activity |

[a] Ts, temperature sensitive.

## 3.4.2. Regulation of fatty acid biosynthesis

The entire pathway is regulated by a slew of mechanisms. Pathway enzymes are feedback-inhibited by long-chain (>12 carbon) acyl-ACPs, and high level expression of acyl-ACP thioesterases, which convert acyl-ACPs to fatty acids, can relieve this inhibition. There are also implications of associated induction of fatty acid degradative pathways with fatty acid overproduction. In general, however, regulation of fatty acid biosynthesis is most affected at an early step in the process, and several pieces of evidence support this [78, 79]. Measurement of acyl-ACP pool size and composition has shown that it is primarily composed of short-chain acyl-ACPs. However, long-chain acyl-ACPs accumulate in vivo when utilization through glycerolphosphate acyltransferase is blocked [78].

As acetyl-CoA carboxylase catalyzes the most important step in fatty acid biosynthesis, it is important to review its regulatory mechanisms. The pathways involved

are seen in Figure 22. The two half-reactions are catalyzed by different subunits of the enzyme, which is wholly composed of four different subunits (designated *accA, accB, accC,* and *accD*). Unfortunately, regulation of *acc* is unclear. The *accB* and *accC* genes are cotranscribed from a promoter far upstream of *accB*. Transcription of all four *acc* genes is related to growth rate - that is, rate of transcription increases with the rate of growth [80]. However, further uncertainty is introduced as the *accBC* operon seems to be regulated by a different mechanism than that which governs the *accA* and *accD* genes [80]. Acetyl-CoA carboxylase is also inhibited by long-chain acyl-ACPs (along with FabI and FabH).[81] The effect of fatty acid production from over-expression of ACC will be detailed later, though in short has resulted in increases in fatty acid production.

The balance between beta oxidation and fatty acid synthesis is mediated by two transcriptional regulators encoded by *fadR* and *fabR*. FadR is involved in the negative regulation of beta oxidation (fatty acid degradation) and in positive regulation of biosynthesis of fatty acids. Specifically, it transcriptionally activates *fabA* and *fabB* and represses the beta oxidation pathway in the absence of long-chain acyl-CoAs.[16] In contrast, FabR inactivates *fabA* and *fabB*.

Over-expression of the β-ketoacyl-acyl ACP synthetases have resulted in inhibition of growth in *E. coli*. Over-expression of FabB has resulted in relatively moderate inhibition, though over-expression of FabF was shown to be extremely toxic to the cells.[82] The effect was found to be the cessation of phospholipid synthesis due to the blockage of fatty acid synthesis. Malonyl-CoA levels increased to about 40% of the total cellular pool of CoA from 0.5% without over-production. Over-production of FabD

partially relieved the toxicity, consistent with the theory that FabF over-production blocks the conversion of malonyl-CoA to malonyl-ACP.

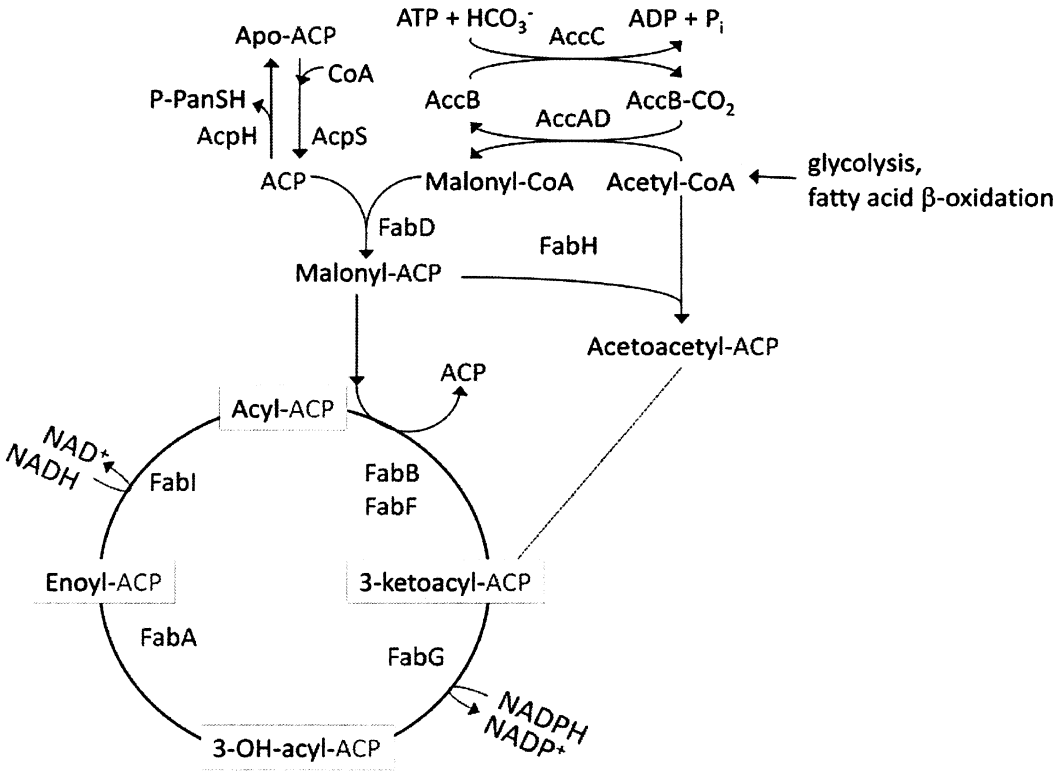



**Figure 23. Overall fatty acid synthesis pathway with relevant enzymes labeled.**

The expression of FabH, FabD, and FabG is coordinated by the *fabH* promoter. During stringent response, phospholipid synthesis is inhibited. It was found that the *fabH* promoter was repressed during stringent response, resulting in the inhibition of phospholipid synthesis and *fabDGH*.[83]

*FabA* and *fabZ* are two genes encoding beta-hydroxyacyl-ACP dehydratases. FabZ efficiently dehydrates the short chain beta-hydroxyacyl-ACPs and long chain

saturated and unsaturated beta-hydroxyacyl-ACPs, while FabA exhibited more activity on intermediate chain length beta-hydroxyacyl-ACPs but still possessed high activity on those that were short and long-chain.[84] FabA has been over-expressed 10-fold in *E. coli*, which resulted in significant increases in saturated fatty acids in phospholipids but no increase in unsaturated fatty acids, indicating that the level of dehydrase activity in wild-type cells does not limit the rate of unsaturated fatty acid synthesis.[85] Unsaturated versus saturated fatty acid production is known to be regulated at this pathway and is thermally regulated.

### 3.4.3. Over-production of fatty acids

Recent studies have targeted some of the more obvious genotypic alterations that could lead to increased fatty acid production. A study in 2008 introduced the genetic alterations as shown in Figure 24.

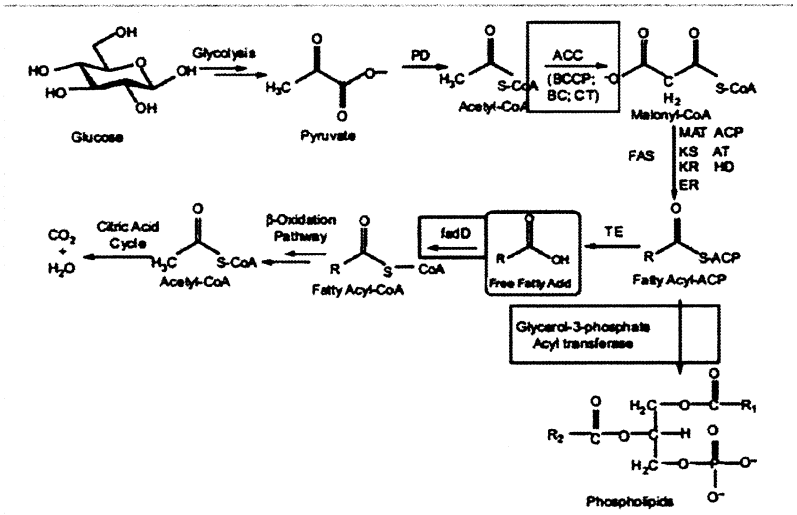



**Figure 24. Lu et. al. 2008 introduced these genetic modifications for several-fold increases in free fatty acid production.**

(Red denotes knocked out pathways, and green denotes over-expressed pathways)

The *fadD* and glycerol-3-phosphate acyl transferase knockouts served to increase the pool of free fatty acids and fatty acyl-ACPs respectively, which resulted in the increase of free fatty acids overall. In addition, over-expression of ACC increased the flux from acetyl-CoA towards malonyl-CoA/ACP, the main building block of fatty acids.

One study published in 2010 reported the development of an engineered *E. coli* strain that was able to overproduce medium-chain length fatty acids through overexpression of acetyl-CoA carboxylase (ACC), elimination of the beta-oxidation pathway through knockout of *fadR* (the first enzyme of the pathway), and introduction of a plant thioesterase known to produce medium chain length fatty acids [31, 86]. The metabolic pathways and the pathway modifications are shown in Figure 25.

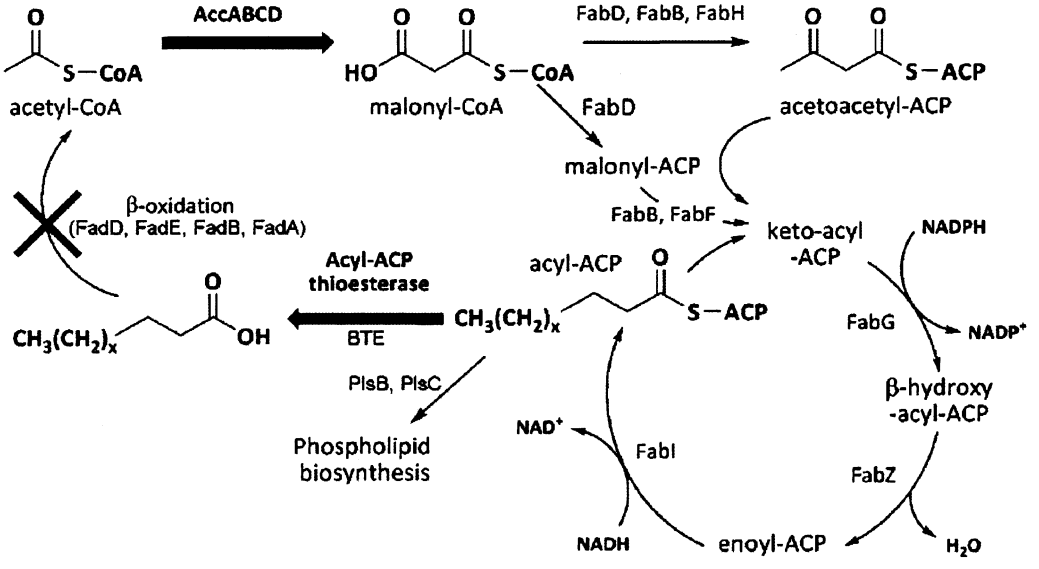



**Figure 25. Pathway modifications introduced in Lennen et. al. 2010 [31].**

The bolded arrows represent overexpressed pathways or heterologous expression (BTE), and X represents disruption by gene deletion.

These two studies focused on redirecting the carbon flux towards the production of free fatty acids, but other studies have accomplished this through increasing co-factor availability. It is known that NADPH is required for fatty acid biosynthesis, and a major supplier of it is thought to be $NADP^+$-dependent malic enzyme [87]. One unpublished study over-expressed ACC, which resulted in a 3-fold increase in cellular lipids compared to the wild-type strain. Over-expression of *E. coli*'s native malic enzyme resulted in a 4-fold increase, and co-expression of the two genes resulted in a 7.5-fold increase.

## 3.5. METHODS FOR EXTRACTION, IDENTIFICATION, AND QUANTIFICATION OF METHYL KETONES

To assess strain performance regarding the production of methyl ketones, it is very important to have a sensitive, precise, and accurate way to measure methyl ketone concentration not only of the culture media supernatant, but also from the intracellular component of the cells. Though there are no published methods for methyl ketone-specific extractions from *E. coli* culture, many methods for fatty acid extraction have been documented and provide a good starting point as fatty acids are similar in structure and properties to methyl ketones.

### 3.5.1. Lysis

If a product is produced intracellularly and not secreted by the cell, or if the product is to be removed from plant, fungal, or animal tissue, then it is necessary to remove the product by lysis, or breaking of the cells or tissues. There are a wide variety of procedures to do so, broadly classified as "chemical" or "physical" methods. The appropriate one chosen based on the nature of the product as well as the nature of the cell

or tissue. In the case of methyl ketones, due to their hydrophobic nature, it is quite possible for them to remain in the cytoplasm as an intracellular product.

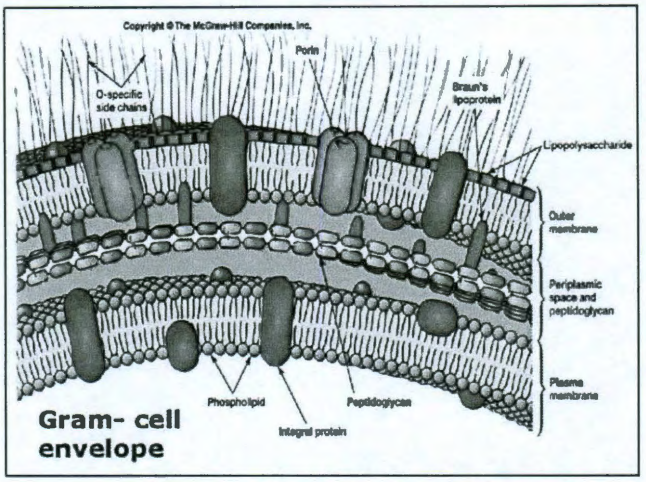



**Figure 26. Gram-cell envelope structure.**

(source: Microbiology, with Diseases by Body System by Bauman, 2nd edition)

Typically, bacterial cell envelopes consist of an inner plasma membrane, a peptidoglycan cell wall, and an outer membrane as shown in Figure 26. Bacterial cells with a thick cell wall are called "Gram positive", whereas those with thin cell walls are "Gram negative" (the names are due to a staining procedure of the peptidoglycan cell wall). For example, *E. coli* is a Gram-negative bacterium. In general, the envelope of the bacterial cell is very resistant to breaking. The membranes themselves are composed of phospholipid bilayers, as shown in Figure 27. The stiffness of the membrane is conferred by the amount of cholesterol and other steroids that happen to exist in the non-polar layer, whereas flexibility is conferred by amphiphilic proteins wedged partially or wholly through the membrane.

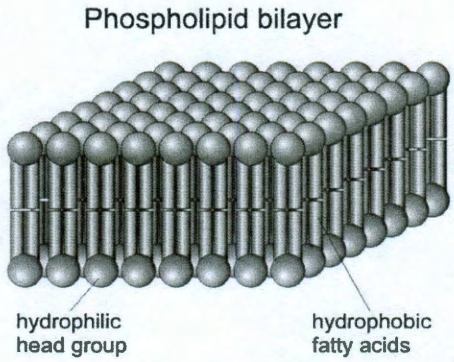



**Figure 27. Phospholipid bilayer structure.**

(source: Forschungzentrum Institute of Structural Biology, http://www.fz-juelich.de/isb/isb-1/datapool/page/28/Figure1-500.jpg)

The bacterial cell wall protects the cell from physical as well as osmotic stress. However, depending on the concentrations of solutes outside of the cell relative to the inside, water will either be drawn out or drawn into the cytoplasm. When too much water is drawn in, it may break the cell wall, a condition that is beneficial for harvesting the retained products.

As aforementioned, various methods exist for the breakage of cells, but each are best suited for different applications. They differ not only their principle mode of action, but also in the stress they impart on the cells/tissues, their cost, and their scale.

Though every cell has an osmotic gradient, drastic changes to extracellular concentrations of solutes can have disastrous impacts on cells that do not have walls, such as animal cells. If transmembrane osmotic pressure is related to the solute concentration inside and outside the cell, van't Hoff's law can be used to estimate this osmotic pressure:

$$\pi = RT(C_i - C_o)$$

where pi is the osmotic pressure, R is the gas constant, T is absolute temperature in Kelvin, and Ci-Co is the difference between solute molarity inside and outside the cell.

Unfortunately, bacterial cells have walls which protect them against osmotic lysis. However, there are other methods to destroy this wall.

Enzymatic lysis can be achieved with the addition of the enzyme lysozyme, which is readily available commercially. However, these are not suitable for applications that require the harvesting of proteins.

Detergents are another way to lyse cells, though they are limited to animal cells in application as they act on plasma membranes. Typically, detergents such as Triton X-100 are used, though it is mainly a trial and error process as different proteins react differently depending on the detergent used.

Finally, and most relevant, is the use of solvents for lysis of cells. An addition of an organic solvent, for example, can dissolve membranes which can be of use after mechanical breaking of cell walls. Benzene, the organic solvent used in this study, is known also to permeabilize cell walls [88].

### 3.5.2. Extraction

Methyl ketones are in general organic liquids that are insoluble in water but quite soluble in common organic solvents such as ethanol, chloroform, benzene, and acetone. In this scenario, liquid-liquid organic extraction is the most suitable tool for extracting methyl ketones. Liquid-liquid extraction is a mass transfer process wherein a liquid solution (feed) is contacted with an immiscible liquid (solvent) that exhibits higher affinity towards the desired solute in the feed. The end result is two streams: the extract, which is the solvent containing the extracted solute, and the raffinate, which is the residual feed minus the extracted solute. However, operating conditions such as solvent

type, contact time, and mixing method need to be carefully evaluated and optimized for a given feed and solute.

Currently, most of the scientific community uses a rapid and reproducible method developed by Bligh and Dyer in 1959 for the extraction of fatty acids from fish tissue [89]. In this method, fatty acids are primarily extracted from tissue using a chloroform-methanol solution. The chloroform-methanol solution, when added to tissue (primarily composed of water), results in a homogenous, monophasic system, which they hypothesized correctly to give optimum lipid extraction. The subsequent addition of water and/or chloroform turns the mixture into a biphasic system, from which the organic layer now rich in lipids can be taken. To be able to analyze these lipids in a gas chromatograph, the chloroform-methanol extract is evaporated away, the leftover lipids acidified to produce fatty acid methyl esters, and finally extracted again using hexane. However, these last derivatization steps do not apply to methyl ketones, as they can be directly analyzed in a gas chromatograph.

# 4. MATERIALS AND METHODS

## 4.1. STRAINS, PLASMIDS, AND GENETIC METHODS

Table 4 describes plasmids used in this study, which express different combinations of *shmks2, slmks2,* and *mks1* on plasmids with copy-numbers spanning from a low of ~5 (pZS), to a moderate level of ~15-20 (pTrcHis2A), to a high level of ~300-500 (pCR2.1-TOPO). The pZS plasmid is a derivative of pZA31-luc, part of the pZ family of plasmids [90-92]. The pTrcHis2A and pCR2.1-TOPO plasmids are supplied by Invitrogen (Carlsbad, California). The genes *shmks2* and *slmks2* are from *Solanum habrochaites* and *S. lycopersicum* respectively, and are 3-ketoacyl-ACP thioesterases, members of the α/β hydrolase super-family. The gene *mks1* is from *S. habrochaites*, and is thought to be a beta-decarboxylase that also exhibits some thioesterase activity. All three genes were amplified from expression vectors supplied by Dr. Pichersky's lab carrying full ORFs of the three genes [21]. The gene *adhF1* was amplified from pABE supplied by Dr. Bornscheuer lab [73].

The genes were cloned into the plasmids using the Clontech (Mountain View, CA) In-Fusion cloning framework, the general scheme of which is shown in Figure 28. The method is designed to join multiple pieces of DNA which have 15 bp's of homology at their linear ends, and is mainly used for the cloning of PCR products linearized vectors without the need for restriction enzyme cleavage of the insert and without the use of ligase or blunt-end polishing. The proprietary enzyme promotes single-strand annealing reactions and can assemble DNA molecules that share short sequence homologies at their ends. Exonuclease activity of the proprietary enzyme cleaves nucleotides from 3' ends,

exposing any overlapping sequence, which become free to anneal and form non-covalently joined molecules that are repaired within the target *E. coli* strain.

In case of the single-expression plasmids, a cloning site was chosen within the multi-cloning site upstream of the promoters and ribosome binding sites as the insertion site for the gene of interest, and the plasmid was linearized. For pZS, KpnI and MluI were used for linearizing the plasmid. For pTrcHis2A, KpnI and SalI were used, and for pCR, KpnI and EcoRI were used. To prepare the cloning inserts, Primer3 (http://frodo.wi.mit.edu/primer3/) was used to design gene-specific primers flanking the gene of interest. Once the primer sequences were found, a 15 bp sequence homologous to the 3' ends of the linearized plasmid were added as shown in the first step of Figure 28. Following PCR amplification of the inserts, they were subsequently run on agarose gels to confirm successful PCR amplification and gel-purified. The inserts were then incubated along with the linearized vector of choice along with the In-Fusion proprietary buffers and enzymes for the cloning reaction. The resultant nicked plasmids were then transformed via heat shock into Clontech's Fusion Blue competent cells, an *E. coli* K-13 strain that provides high transformation efficiency and contains *recA* and *endA* mutations, making it a good host for high yields of plasmid DNA. After transformation, the cells were plated on LB plates with appropriate antibiotics for screening of transformed cells for plasmids. Proper cloning of the inserts was checked through colony-PCR.

In the case of the double-expression plasmids, the procedure was essentially the same except the primers used for the amplification of the second gene (*mks1*) involved an addition of a ribosome binding site consensus sequence (GAGGAG) along with 6 dummy base pairs between the 15 bp homology to the vector and the gene-specific priming

sequence [93]. The PCR insert derived from these primer sets were cloned downstream of the first gene's ORF (*shmks2* or *slmks2*). In this way, an operon would be constructed where both genes would be transcribed and controlled under the same promoter, with translation occurring for the two genes on one transcript due to the ribosome binding sites. A map of this is shown in Figure 29.

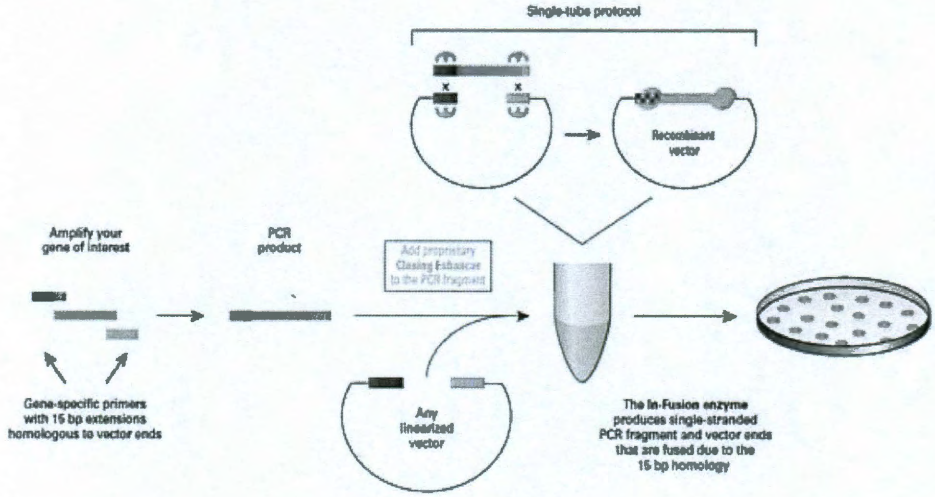



**Figure 28. Clontech's In-Fusion protocol.**

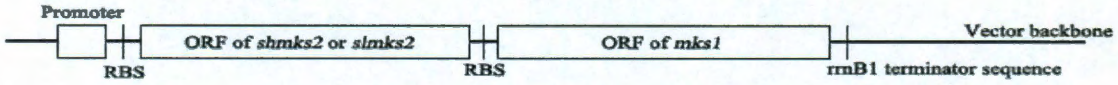



**Figure 29. Map of double expression operon.**

Afterwards, plasmids miniprepped from the Fusion Blue competent cells were transformed into MG1655 *E. coli* cells by electroporation and selected on LB plates with appropriate antibiotics (chloramphenicol for pZS series, ampicillin for pTrcHis2A series

and pCR series). Gene expression was induced with 100 ng/mL anhydrous tetracycline for the pZS series plasmids, and 100 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) for the pTrcHis2A and pCR series plasmids. Plasmids miniprepped from the final strains were verified through both restriction analysis and sequencing of the operon regions.

Table 5 contains the backgrounds, genotypes, and sources of the strains of *E. coli* used in this study. All strains used are derivatives of *Escherichia coli* K-12 strain MG1655 (F- λ- *ilvG- rfb*-50 *rph*-1), obtained from the University of Wisconsin *E. coli* Genome Project (www.genome.wisc.edu) [94]. Mutant alleles were moved into this background via standard P1 transduction as described in Blankschien et. al.[95]

**Table 4. Plasmid list.**

| Plasmid | Relevant information | Source |
|---|---|---|
| pZSblank | Blank plasmid created by removing *C. freundii dhaKL* fragment from pZSKLcf and self-ligating the plasmid (tetR, oriR SC101*, cat) | Yazdani and Gonzalez 2008 [92] |
| pZS-shmks2 | *Solanum habrochaites shmks2* gene under control of $P_{Ltet0-1}$ (tetR, oriR SC101*, cat) | This study |
| pZS-slmks2 | *S. lycopersicum slmks2* gene under control of $P_{LtetO-1}$ (tetR, oriR SC101*, cat) | This study |
| pZS-mks1 | *S. habrochaites mks1* gene under control of $P_{LtetO-1}$ (tetR, oriR SC101*, cat) | This study |
| pTrcHis2A-shmks2 | *S. habrochaites shmks2* gene under control of $P_{lacZ}$ (lacR, oriR BR322) | This study |
| pTrcHis2A-slmks2 | *S. lycopersicum slmks2* gene under control of $P_{lacZ}$ (lacR, oriR BR322) | This study |
| pTrcHis2A-mks1 | *S. habrochaites mks1* gene under control of $P_{lacZ}$ (lacR, oriR BR322) | This study |
| pTrcHis2A-shmks2-mks1 | *S. habrochaites shmks2* and *S. habrochaites mks1* genes under control of $P_{lacZ}$ (lacR, oriR BR322) | This study |
| pTrcHis2A-slmks2-mks1 | *S. lycopersicum slmks2* and *S. habrochaites mks1*genes under control of $P_{lacZ}$ (lacR, oriR BR322) | This study |
| pCR-shmks2 | *S. habrochaites shmks2* gene under control of $P_{lacZ}$ (oriR pUC) | This study |
| pCR-slmks2 | *Solanum lycopersicum slmks2* gene under control of $P_{lacZ}$ (oriR pUC) | This study |
| pCR-shmks2-mks1 | *S. habrochaites shmks2* and *S. habrochaites mks1* genes under control of $P_{lacZ}$ (oriR pUC) | This study |
| pCR-slmks2-mks1 | *S. lycopersicum slmks2* and *S. habrochaites mks1*genes under control of $P_{lacZ}$ (oriR pUC) | This study |
| pZS-adhF1 | *Pseudomonas fluorescens adhF1* gene under control of $P_{Ltet0-1}$ (tetR, oriR SC101*, cat) | This study |
| pTrcHis2A-adhF1 | *Pseudomonas fluorescens adhF1* gene under control of $P_{lacZ}$ (lacR, oriR BR322) | This study |

**Table 5. Strain list**

| Plasmid | Relevant information | Source |
|---|---|---|
| MG1655 | F-l-ilvG-rfb-50 rph-1 | Kang et. al. (2004)[94] |
| Δ*adhE* | MG1655 Δ*adhE*::FRT-Kan-FRT | This study |
| Δ*ldhA* | MG1655 Δ*ldhA*::FRT-Kan-FRT | This study |
| Δ*poxB* | MG1655 Δ*poxB*::FRT-Kan-FRT | This study |
| Δ*pta* | MG1655 Δ*pta*::FRT-Kan-FRT | This study |
| Δ*adhE* Δ*pta* | MG1655 Δ*adhE*::FRT Δ*pta*::FRT-Kan-FRT | This study |
| Δ*ldhA* Δ*poxB* | MG1655 Δ*ldhA*::FRT Δ*poxB*::FRT-Kan-FRT | This study |
| Δ*ldhA* Δ*poxB* Δ*pta* | MG1655 Δ*ldhA*::FRT Δ*poxB*::FRT Δ*pta*::FRT | This study |
| Δ*adhE* Δ*ldhA* Δ*poxB* Δ*pta* | MG1655 Δa*dhE*::FRT Δ*ldhA*::FRT Δ*poxB*::FRT Δ*pta*::FRT | This study |

## 4.2. Culture Media and Cultivation Conditions

### 4.2.1. Culture Media

All work was done at 37°C with either LB medium or MOPS minimal medium (below), supplemented, when required, with sodium citrate (5-20 mM, for phage P1 transductions), ampicillin (50 ug/mL), chloramphenical (34 ug/mL), glucose (10 g/L unless otherwise stated), IPTG (0.1 mM), and anhydrotetracycline (100 ng/mL). The minimal media designed by Neidhardt (MOPS) was supplemented with 1.32 mM $Na_2HPO_4$ in place of $K_2HPO_4$, 5 mM $(NH_4)_2SO_4$, and 30 mM $NH_4Cl$ unless otherwise

stated [96]. All chemicals for culture media were from Fisher Scientific (Pittsburgh, PA) and Sigma-Aldrich Co. (St. Louis, MO).

Long-term storage of strains was accomplished by preparing glycerol stocks as described in *Current Protocols in Molecular Biology* [97]. Glycerol stock media was prepared as a filter-sterilized 2× solution containing 65% (v/v) glycerol, 0.1 M MgSO4, and 0.025 M Tris-HCl pH 8. Equal amounts of this solution and media containing exponentially growing cultures (OD ~0.5) were combined and subsequently stored at -80 °C for future use. These stocks were used to streak media plates containing 10 g/L NaCl, 5 g/L yeast extract, 10 g/L tryptone and 1.5% agar. When required, antibiotics were included in the following concentrations: 100 μg/mL ampicillin, 34 μg/mL chloramphenicol, 50 μg/mL kanamycin, and 12.5 μg/mL tetracycline.

. Calcium carbonate, which is biologically inert, was added at a concentration of 50 g/L for pH control of flask cultures, unless otherwise stated.

### 4.2.2. Cultivation conditions - Shake flask experiments

Prior to inoculating starter cultures, frozen stocks (stored as glycerol stocks at -80 °C) were streaked onto LB plates (with appropriate antibiotics if required) and incubated overnight at 37 °C. A single colony was used to inoculate the starter culture flasks.

Starter cultures were cultivated in LB with appropriate antibiotics at a 10 mL working volume in 25 mL Erlenmeyer flasks. The operating conditions were 37°C at 200 RPM in a New Brunswick Scientific (Edison, NJ) benchtop incubator. Starter cultures were taken out during exponential phase (OD ~0.5), and depending on the working volume of the experimental flask culture to be inoculated, a certain volume was pelleted

down at 6000 g for 4 min to inoculate the experimental flask cultures to an OD of 0.05. If the experimental flask culture media was minimal media (MOPS supplemented as stated above), the pellets were washed twice with the minimal media (centrifuging at 6000 g for 4 min each time) to reduce the increased lag phase the cells may experience when switching to different media.

The experimental flask cultures were cultivated in either minimal media or LB-Miller with appropriate antibiotics at various working volumes in 125 mL Erlenmeyer flasks. The operating conditions were 37°C at 200 RPM in a New Brunswick (Edison, NJ) benchtop incubator.

### 4.2.3. Cultivation conditions - Fermentor experiments

Prior to inoculating starter cultures, frozen stocks (stored as glycerol stocks at -80 °C) were streaked onto LB plates (with appropriate antibiotics if required) and incubated overnight at 37 °C. A single colony was used to inoculate the starter culture flasks.

Pre-starter cultures were cultivated in LB-Miller with appropriate antibiotics at a 10 mL working volume in 25 mL Erlenmeyer flasks. The operating conditions were 37°C at 200 RPM in a New Brunswick Scientific benchtop incubator. Pre-starter cultures were taken out during exponential phase (OD ~0.5), and depending on the working volume of the experimental flask culture to be inoculated, a certain volume was pelleted down at 6000 g for 4 min to inoculate the starter cultures to an OD of 0.05. Pre-starter culture was required for reasonable growth time in the primary starter culture grown in MOPS minimal media, described below.

The starter cultures were cultivated in MOPS minimal media as described previously with 10 g/L glucose and appropriate antibiotics. The operating conditions were 37°C at 200 RPM in a New Brunswick benchtop incubator. Starter cultures were taken out during exponential phase (OD ~0.5), and depending on the working volume of the experimental flask culture to be inoculated, a certain volume was pelleted down at 6000 g for 4 min to inoculate the fermentor cultures to an OD of 0.05. MOPS minimal media was used for the primary starter culture to decrease the lag time between inoculation and exponential phase in the fermentor.

The fermentor cultures were cultivated in a Infors HT (Bottmingen, Switzerland) Sixfors bench-top six bioreactor setup. Each vessel was fitted with an impeller, dissolved oxygen probe, pH probe, acid and base inlets, condenser with headspace outlet, gas inlet (submerged for sparging of media), temperature probe, and submerged outlet tube for culture collection. Externally, a magnetic stirrer and thermo-electric block provided impeller RPM and temperature control, respectively. A .45 micron filter was attached to the headspace outlet from the condenser, and a .2 micron filter was attached between the gas inlet and the gas source. The controlled variables included pH, dissolved oxygen, and temperature, and these variables were constantly measured through the Sixfors computer via sensor attachments. pH was adjusted with either 1.5 M sulfuric acid or 3 M sodium hydroxide. Dissolved oxygen was controlled by RPM adjustment between 200-800 via the Sixfors microcontroller. From inoculation, air was supplied at a rate of approximately .5 L/min by an external air pump. Over the course of fermentation, maintenance of dissolved oxygen levels required the sparging of pure oxygen (Oxygen, compressed, ultra high purity, Matheson Tri-Gas). Fermentor conditions were carefully

monitored over the course of the entire fermentation to see when such a switch from air to oxygen was necessary. The oxygen was provided at a rate of between 0.1 L/min to 0.5 L/min, manually adjusted during the fermentation to keep dissolved oxygen levels constant.

## 4.3. OPTIMIZATION OF EXTRACTION METHODS

### 4.3.1. Lysis methods

The mechanical breakage of cells was conducted using the Disruptor Genie (Scientific Industries Inc., Bohemia, New York) with 0.1 mm diameter glass beads (Scientific Industries Inc., Bohemia, New York) designed for bacterial cells. Assessment of the optimal operating conditions for the ball mill method was done by varying time of disruption (power level is constant and cannot be changed by the user).

Sonication was assessed using the Branson Sonifier 450 (Emerson Industrial Automation, Danbury, CT) with the tip attachment. As the equipment literature stated, the area-of-effect of the sonicator is shaped like a cone projecting from the botton of the tip. Thus, the tip was inserted approximately 1 cm into the samples, the minimum depth recommended by the manufacturer (twice the diameter of the tip). The equipment literature stated typical complete bacterial cell breakage will occur at 45 seconds of sonication. With cell culture volume and power setting fixed, time of sonication was varied from 0, 45, 90, and 180 seconds for various ODs. Duty cycle (the portion of each second during which the sonication occurs) was set to 30% so as to give the sample time to cool.

### 4.3.2. Extraction methods

Benzene, n-hexane, and chloroform were obtained from Sigma-Aldrich Co. (St. Louis, MO). Significant operating conditions such as vortex time and solvent/feed volumes were varied, keeping all factors regarding the feed constant across all the variations. Vortex time was varied from 0, 30, 60, 90, 120, and 150 seconds. Solvent/feed volume was determined using an iterative process, steadily increasing the culture volume while keeping solvent volume constant at 1 mL such that the peak reading becomes reliable on the GC-FID. All other equipment and methods regarding the extraction can be found in section 3.4.2.

## 4.4. Analytical Methods

### 4.4.1. Optical density, pH, and HPLC analysis

Optical density was measured at a wavelength of 550 nm on a Genesys spectrophotometer and used as an estimate of cell mass (1 $OD_{550}$ = 0.34 g dry weight/L). When required, the pH of samples were measured using a Denver UB-10 pH meter (Denver Instruments, Denver, CO). After centrifugation of the sample, the supernatant was stored at -20 °C for future HPLC analysis. Quantification of the concentration of glucose and other compounds was accomplished with ion-exclusion HPLC by injection of 10 μL of sample through a Shimadzu Prominence SIL 20 system (Shimadzu Scientific Instruments Inc., Columbia, MD) equipped with an HPX-87H organic acid column (Bio-Rad, Hercules, CA). Operating conditions were optimized for adequate peak separation (30 mM H2SO4 in mobile phase at 0.3 mL/min, column temperature 42 °C), previously determined by Dharmadi and Gonzalez [98].

### 4.4.2. Extraction of methyl ketones in culture

The whole process of extraction, identification, and quantification starts with at least 15 mL of culture for accurate analysis. For example, to run an analysis of a flask culture with a working volume of 10 mL, another replicate must have been run in parallel to meet the 15 mL culture requirement for analysis. This is to adequately acquire the 10 mL requirement for extraction and the 2 mL required for supernatant storage for future HPLC analysis.

Extraction began with splitting 10 mL of culture into two 5 mL aliquots each in 14 mL round-bottom polypropylene Falcon tubes, which are wide to prevent the sonication tip from hitting the walls. Each Falcon tube holding 5 mL of culture was placed in an ice bath and set in a Branson Sonifier 450 (Emerson Industrial Automation, Danbury, CT) such that the sonicator tip was approximately 1 cm deep and centered. The sample was subsequently sonicated for 5 minutes at a duty cycle of 30% and a power setting of 1.

Once sonication was complete for both 5 mL aliquots, they were combined into a 15 mL Corning CentriStar centrifuge tube (Corning Inc., Lowell, MA) for the liquid-liquid extraction. 1 mL of benzene (deuterated when preparing samples for NMR analysis) was added, and the solution vortexed for 2 minutes. Afterward, the solution was placed on a Thermo Scientific Barnstead Labquake rotisserie (Thermo Fisher Scientific Inc., Asheville, NC) for 1 hour to maximize methyl ketone transfer and to allow the aqueous and organic droplets to settle. Finally, the solution was centrifuged for 15 minutes at 12,000 g to separate the phases, and the organic phase recovered.

The recovered extract was placed in a 2-mL screwtop PTFE glass vial and stored at -20°C for future GC-MS/FID or NMR analysis. The overall flow of this procedure is shown in Figure 30. For preparation of supernatant feed (as opposed to the sonicated whole culture feed), the culture liquid was spun down at 12,000 g for 5 minutes in 15- or 50- mL Corning CentriStar centrifuge tubes (Corning Inc., Lowell, MA), and the appropriate volume of supernatant recovered.

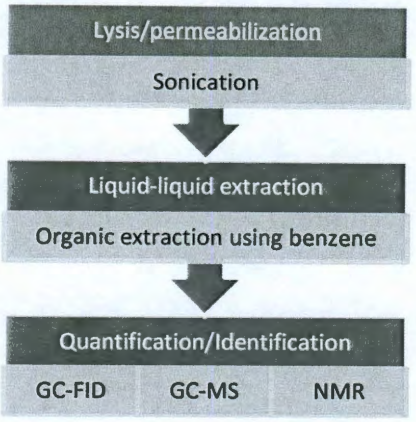



**Figure 30. Extraction, identification, and quantification protocol at a glance.**

### 4.4.3. Identification of methyl ketones

Identification of methyl ketones consisted of two approaches: one from GC-MS, and another through one-dimensional proton NMR. The NMR results provided structural verification by confirming the presence of the first 5 methyl ketone carbons, and the GC-MS provided total mass information which revealed the chain lengths of the methyl ketones present in the sample.

GC-MS was conducted by injecting 1 μL of sample through a Shimadzu GCMS-QP2010S (Shimadzu Scientific Instruments Inc., Columbia, MD) equipped with an

Agilent DB-5 (Agilent Technologies, Foster City, CA) (5%-Phenyl)-methylpolysiloxane, non-polar, high temperature column. The carrier gas was helium, and the injector was set to hold at 250°C for 0.5 min, ramp to 280C at 50C/min, and hold for 1 min. The column oven was set to initialize at 50°C, hold for 5 min, then ramp to 275 at a rate of 3.75°C/min. Total time per sample was approx 65 min. Identification of compounds was conducted using the similarity search algorithm through built-in libraries of Shimadzu's GCMS Solution software. Peaks that came up as methyl ketones in the similarity search were further confirmed by comparing their retention times with those from methyl ketone standards.

NMR was conducted by aliquoting approximately 600 μL of extract (using deuterated benzene) in a glass NMR tube for reliable shimming of the sample, and a Varian Inova 500 (500 MHz NMR) equipped with a HCNP quad probe (Varian Inc., Palo Alto, CA) was used to obtain the 1-D proton spectra. The following parameters will be utilized: 25°C sample temperature, 8,000 Hz sweep width, 2.8 s acquisition time, 256 acquisitions, 6.3 μs pulse width, 1.2 s pulse repletion delay, and presaturation for 2 s. Prior to running the samples, methyl ketone standards at a concentration of 50 mg/L were used to identify the signature peaks and their associated J-coupling and chemical shift values.

### 4.4.4. Quantification of methyl ketones

Quantification of methyl ketones was primarily done by injecting 1 μL of sample through GC-FID analysis using a Varian CP-3800 GC (Varian Inc., Palo Alto, CA) with an FID back-end and equipped with a Varian VF-5ht 5% phenyl methyl dimethylpolysiloxane, non-polar, high temperature column. Prior to sample analysis but

following temperature program optimization, standards of 3 methyl ketones (2-nonanone, 2-undecanone, 2-tridecanone) with concentrations of 5, 10, 20, 50, and 100 mg/L were injected. The peak areas were calculated through the Varian GC software and calibration curves made to relate peak area to methyl ketone concentration in benzene. The carrier gas was helium, and the injector was set to hold at 250°C for 0.5 min, ramp to 280C at 50C/min, and hold for 1 min. The column oven was set to initialize at 50°C, hold for 5 min, then ramp to 275 at a rate of 3.75°C/min. Total time per sample was approx 65 min.

## 4.5. CELL EXTRACT PREPARATION AND SDS-PAGE GEL ANALYSIS

### 4.5.1. *E. coli* extract preparation

Prior to inoculating starter culture, frozen stocks (stored as glycerol stocks at -80 °C) were streaked onto LB plates (with appropriate antibiotics if required) and incubated overnight at 37 °C. A single colony was used to inoculate a 1-5 mL starter culture in LB media in a 15 mL conical Falcon tube overnight. A 25 mL culture of LB media containing 0.1 mM IPTG or 100 ng/mL anhydrotetracycline and appropriate antibiotic was then inoculated with 50 uL of the overnight starter culture and incubated for >2 hours until mid-log growth ($OD_{550}$ of 0.5-1).

Approximately 0.15 $OD_{550}$ units of culture was spun down at 13,000 g for 1 minute. The supernatant was removed and the pellet resuspended in 20 uL of 1X NuPAGE LDS (Invitrogen) sample buffer with NuPAGE reducing agent (Invitrogen). The samples were then heated to 95°C for 15 minutes.

### 4.5.2. SDS-PAGE gel running conditions

The samples as prepared in the previous section were loaded into a pre-cast NuPAGE® Novex 12% Bis-Tris 10-well gel (Invitrogen) secured in a Xcell SureLock Mini Cell (Invitrogen). The upper buffer chamber was filled with approximately 200 mL of 1X NuPAGE SDS running buffer (Invitrogen) supplemented with 500 uL NuPAGE antioxidant. The lower buffer chamber was filled with approximately 600 mL of 1X NuPAGE SDS running buffer. Novex Sharp Unstained Protein standard was used as the protein ladder. A 200 V electric field was applied to enable separation of the proteins across the gel by a Fisher Scientific FB300 power supply (Thermo Fisher Scientific), until the blue marker dye migrated to approximately 1 mm from the bottom of the gel.

### 4.5.3. Gel staining

Once the electrophoresis completed, the gel was washed three times for 5 minutes on an orbital shaker to remove SDS and buffer salts using deionized water. Afterwards, approximately 20 mL of SimplyBlue Safestain (Invitrogen) was decanted such that the gel was submerged completely, and shaken on an orbital shaker overnight at room temperature. De-staining was done by adding 100 mL deionized water and shaking on an orbital shaker at room temperature for 3-5 hours.

# 5. RESULTS AND DISCUSSION

## 5.1. PATHWAY ANALYSIS FROM GLUCOSE TO METHYL KETONES

The knowledge of the two biochemical reactions leading to methyl ketones from polyketide intermediates as described in the literature[21] allowed pathway analysis from glucose to the end product methyl ketone. Figure 31 shows the methyl ketone synthesis reactions attached to the polyketide synthesis cycle and the preceding glycolytic reactions within *E. coli*.

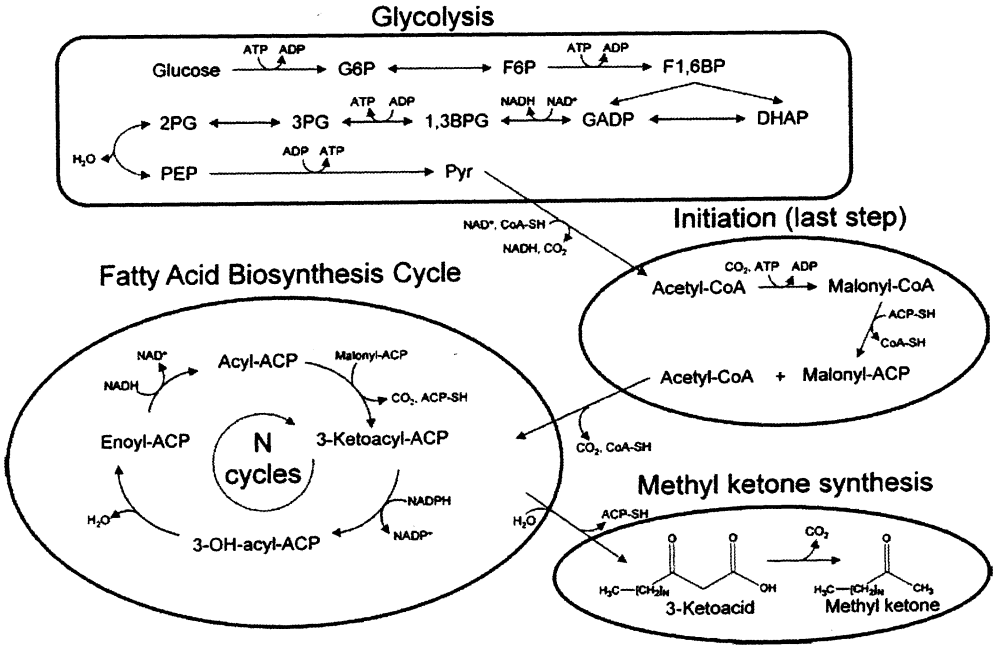



**Figure 31. Overall pathway from glucose to methyl ketones.**

The top portion involves standard glycolysis connecting glucose to pyruvate, the precursor to acetyl-CoA which eventually becomes the indirect building block of fatty acids and methyl ketones. The purple circle outlines the initiation reactions that need to take place for both the first molecule that enters the fatty acid cycle (acetoacetyl-ACP)

and malonyl-ACP, the direct building block that gets added onto acyl-ACP at each turn of the cycle. The blue circle shows the fatty acid synthesis cycle, and the bottom circle outlines the irreversible "termination" reactions that lead out from the fatty acid synthesis cycle at 3-ketoacyl-ACP to form methyl ketones. The stoichiometric balance from glucose to acetyl-CoA (AcCoA), from acetyl-CoA to methyl ketones of various sizes, and then a combined balance from glucose to methyl ketones of various sizes was obtained by tracing the metabolites along the pathway. These balances are shown in Figure 32.

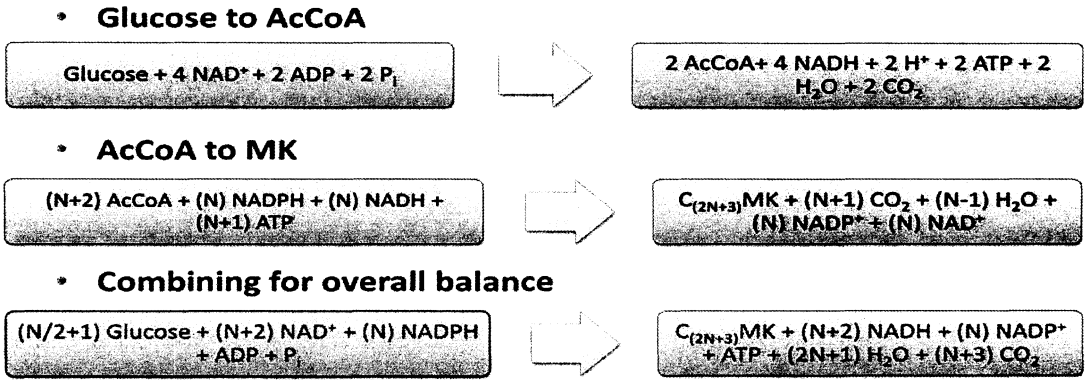



**Figure 32. Balances from glucose to methyl ketones.**

This calculation gives the net change in co-factors and maximum theoretical yields, clues as to what modifications may need to be made for co-factor balance, what environmental conditions may be most appropriate for maximizing product yield, and gives upper bounds for economic purposes. The "N" corresponds to the number cycles the fatty acid synthesis pathway has made in order to create a $C_{2N+3}$ methyl ketone, as shown in Figure 31. This overall balance shows that:

- the process generates ATP independent of the chain length of the resulting methyl ketone, meaning the process to convert glucose to methyl ketones will most likely not significantly inhibit cell growth by competing for ATP with the cell's native anabolic processes.
- the process generates excess reducing equivalents independent of the chain length of the resulting methyl ketone, meaning the conversion of glucose to methyl ketones will require an external electron acceptor such as oxygen.
- the process generates excess NADH but consumes NADPH, which means modifications to increase the pool of NADPH may be beneficial for production of methyl ketones.

## 5.2. ENABLING PRODUCTION OF METHYL KETONES IN *E. COLI*

### 5.2.1. Selection of enzymes

As *E. coli* does not natively produce methyl ketones, the first objective was to graft the two-step methyl ketone synthesis pathway into the wild-type *E. coli* strain, verify the expression of the enzymes, and measure their activities indirectly through identification and quantification of the methyl ketone products. Selection of the appropriate genetic sources for heterologous enzymes is generally difficult due to the unpredictable nature of product expression and activity within a non-native host. However, the possibility of constructing a functional pathway was increased by selecting well-characterized beta-ketoacyl-ACP thioesterases and beta-ketoacid decarboxylases (termed "methyl ketone synthases (MKSs)" by the authors characterizing the enzymes) proven to have some functionality in *E. coli*.[21] Previous studies had shown the presence

of 2-tridecanone (major product) and 2-undecanone (minor product) from heterologous expression of SHMKS2 (beta-ketoacyl-ACP thioesterase from *Solanum habrochaites*) in *E. coli* and similar results with heterologous expression SLMKS2 (beta-ketoacyl-ACP thioesterase from *Solanum lycopersicum*). In-vitro assays of purified MKS1 (beta-ketoacid decarboxylase) by the same authors characterizing the SHMKS2/SLMKS2 showed significantly higher decarboxylase activity of MKS1 compared to SHMKS2, confirming the hypothesis that SHMKS2/SLMKS2 catalyze the first thioester hydrolysis step, and MKS1 catalyzes the second decarboxylation step. The DNA sequences encoding the enzymes were codon-optimized to ensure adequate expression in *E. coli*.

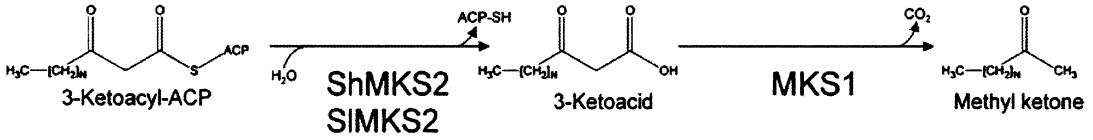



**Figure 33. The newly introduced methyl ketone synthesis reactions and corresponding enzymes.**

## 5.2.2. Construction and evaluation of plasmid performance

Three vectors were used to express the MKSs in *E. coli* at different levels of expression. Plasmid replication and production of heterologous proteins are known to place a physiological burden on cells and may even be toxic to the cells at certain levels of expression, necessitating the need to evaluate the effect of different levels of expression on methyl ketone production in order to determine the optimal balance between cell growth and product formation. The pZS vector was chosen for its low-copy number of ~5 (pSC101 origin), the pTrcHis2A vector for its moderate copy number of ~15-20 (pBR322 origin), and the pCR2.1 vector for its high copy number of 300-500

(pUC origin). Initially, the *shmks2* or *slmks2* gene catalyzing the first step of the methyl ketone synthesis pathway was cloned into the three vectors, as SHMKS2 and SLMKS2 has exhibited the ability to produce methyl ketones by itself without the aid of MKS1,[21] the assumption being that ketoacids produced by the enzyme spontaneously decarboxylate and/or the enzyme has minor decarboxylation activity at a high enough rate for methyl ketone to be produced at a detectable level. The strain background for this initial evaluation of plasmid performance was *E. coli* MG1655, and cultivations were conducted in 10 mL LB-Miller medium at 37°C with appropriate promoter induction performed during culture inoculation. Functionality of the pathways expressed under these various vectors was evaluated by measuring the concentration of methyl ketones produced (extracted through methods outlined in section 3.5).

Unfortunately, by 12h the strain expressing SHMKS2 on pTrcHis2A (MG1655 [pTrcHis2A-shmks2]) was the only one producing a detectable amount of methyl ketones. The strain carrying pZS-shmks2 grew only slightly better than the producing strain but did not produce methyl ketones at a detectable level, which may mean that SHMKS2 thioesterase activity may be so low as to require a higher level of expression to produce a significant amount. The strain MG1655 [pCR2.1-shmks2] grew extremely slowly, taking almost 24h to reach log-phase and even then producing no detectable amounts of methyl ketones, pointing to plasmid instability. One possible cause for this is that the protein SHMKS2 may be significantly toxic to *E. coli* at such high levels. Physiological burden due to plasmid replication is a less likely cause, as many other studies in the literature have used pCR2.1 successfully for overproduction of other proteins. All strains expressing SLMKS2 did not produce detectable amounts of methyl

ketones, possibly due to hindered activity *in vivo* in the non-native environment of *E. coli*. Verification of protein expression and identification of products in the producing strain will be detailed in the next section.

With pTrcHis2A-shmks2 being the most functional construct, the gene *mks1* was cloned in to assemble a single polycistronic operon expressing both *shmks2* and *mks1* in order to fully introduce the non-native methyl ketone synthesis pathway, yielding the plasmid pTrcHis2A-shmks2-mks1. Contrary to expectations, cultivation of the strain carrying this plasmid yielded almost no increase in methyl ketone levels over the strain expressing only SHMKS2. This implies that the thioester hydrolysis catalyzed by SHMKS2 is the rate-limiting step between the two steps of methyl ketone synthesis. Another possibility is that the rate of spontaneous decarboxylation far exceeds the decarboxylase activity of MKS1. However unnecessary MKS1 may seem, the co-expression of *mks1* did not inhibit growth significantly. These results led to the plasmid pTrcHis2A-shmks2-mks1 being chosen as the expression system that conveyed the most functional and complete methyl ketone synthesis pathway, and was the primary plasmid used in further strain characterization studies. As aforementioned, verification of protein expression and identification of products in this strain will be detailed in the next section.

### 5.2.3. Verification of enzyme expression and identification of products

Expression of proteins SHMKS2 and MKS1 were verified by SDS-PAGE gel analysis of whole cell extracts. The bands for SHMKS2 and MKS1 when expressed by the pZS vector were unable to be discerned from the background banding, most likely

due to low expression. Protein expression was not verified for pCR2.1-shmks2 or pCR2.1-shmks2-mk1 due to aforementioned plasmid instability issues.

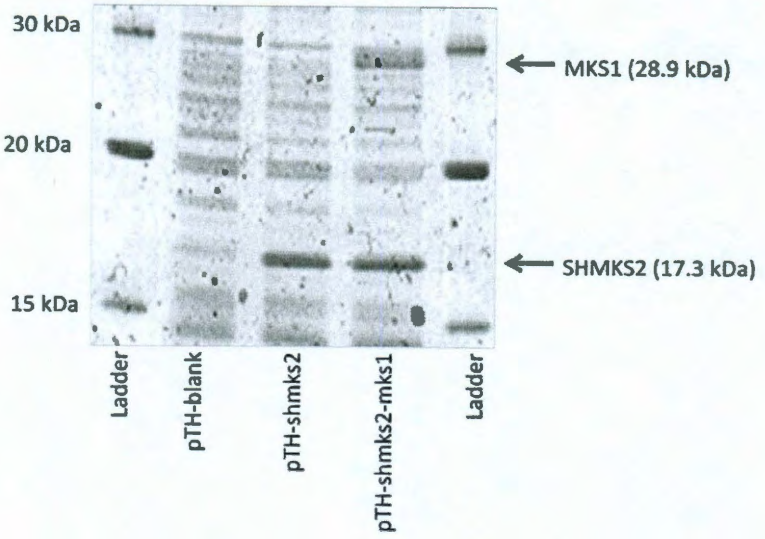



**Figure 34. SDS-PAGE gel analysis for verification of protein expression.**

"pTH" is short for pTrcHis2A. "pTH-blank" denotes the pTrcHis2A vector as supplied by Invitrogen without any cloned genes.

Product identification came in two forms: one-dimensional proton NMR and GC-MS. The following shows the identification of the three methyl ketone products detected as produced by MG1655 [pTrcHis2A-shmks2-mks1]. The extract was derived from a sonicated minimal media liquid culture that was obtained 20 hours after inoculation, using the established techniques as described previously in section 3.5.

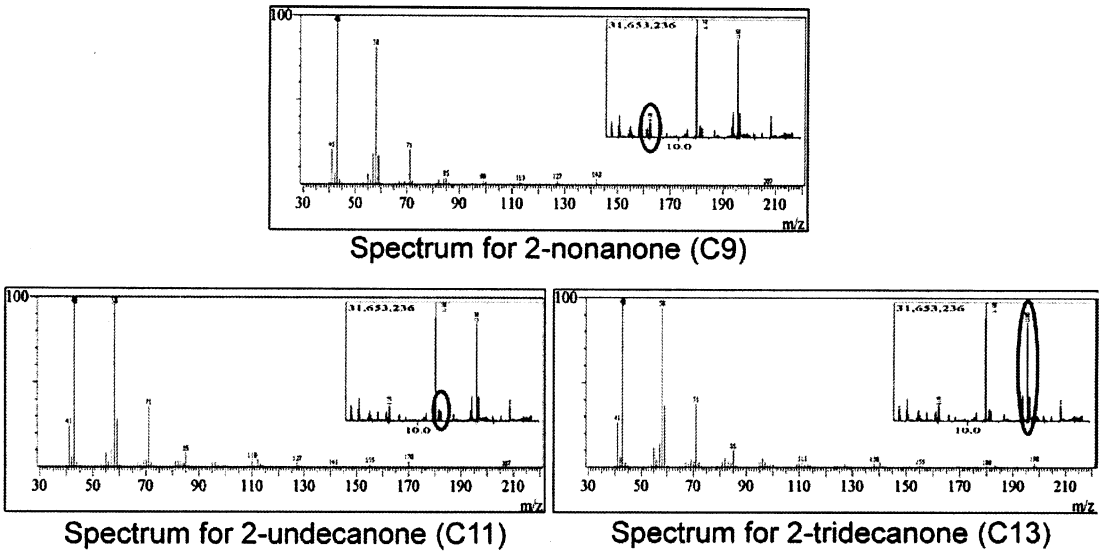



**Figure 35. GC-MS identification of 3 methyl ketones as produced by MG1655 [pTrcHis2A-shmks2-mks1]**

Figure 35 shows the mass spectra associated with the methyl ketone labeled under them, and an inset chromatogram showing the peak from which the spectra was obtained. The calculated concentrations of methyl ketones were in agreement with those calculated from the GC-FID. 2-Tridecanone was the major product by a significant margin, followed by 2-undecanone, and finally 2-nonanone. 2-Nonanone was detected by the GC-MS due to heightened sensitivity of that instrument, but was not ever detected in the GC-FID. Controls using the blank pTrcHis2A plasmid were also analyzed, and no methyl ketones were detected in those samples.

The NMR results also agree with the GC-MS results, as shown in Figure 36. The red boxes indicate where the signature peaks are aligned across the NMR spectra. The H1 corresponds to the chemical shift of the protons on the first carbon (methyl next to the ketone), the main signature peak of a methyl ketone. The H3 corresponds to the chemical shifts of the protons of the third carbon, which again we see in the sample and not in the

control. The inset chemical structure shows the extent to which a chemical structure can be implied by the peaks surrounded in red (the H1 and H3 peaks). Unfortunately, chain length of methyl ketones cannot be determined even in pure methyl ketone samples due to the overlapping peaks in the aliphatic regions around 1.2-1.3 ppm.

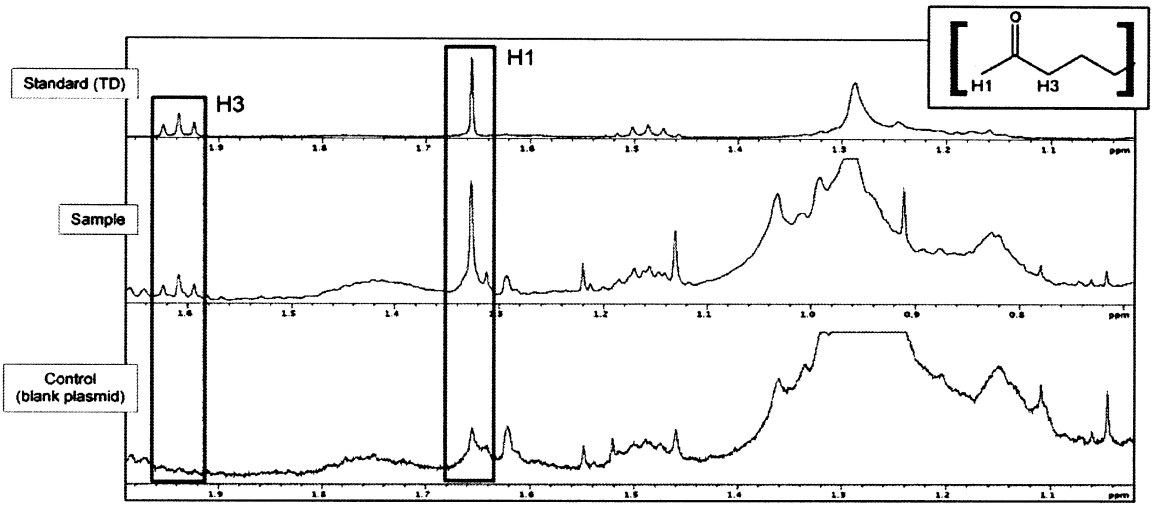



**Figure 36. NMR results comparing standard, sample, and control peaks.**

Subsequent product quantifications were done by integration of GC-FID peaks after the peak identities were confirmed from NMR/GC-MS.

## 5.3. ESTABLISHMENT OF OPTIMAL OPERATING CONDITIONS

From the pathway balances, it is known that the conversion of glucose to methyl ketones generates excess reducing equivalents, thus requiring the presence of an external electron acceptor to achieve redox balance. Though aerobic conditions meet this need, the TCA cycle which is active under these conditions compete for acetyl-CoA with the fatty

acid synthesis cycle, decreasing the yield of methyl ketones from glucose. Microaerobic conditions have been previously utilized to achieve redox balance while still maintaining the ability to synthesize reduced products at high yields.[99] Similarly, it was hypothesized that an optimal microaerobic condition exists between aerobic and anaerobic conditions which will increase yields of the more-reduced methyl ketones from glucose due to inactivation of the competing TCA cycle while still providing enough oxygen for oxidative phosphorylation to close the redox balance.

Using both the MG1655 [pTrcHis2A-shmks2-mks1] and MG1655 [pTrcHis2A-shmks2] strains that exhibited the highest methyl ketone production thus far, the effect of aerobicity was examined roughly by altering shake flask working volume. By increasing the working volume, the surface area available for oxygen transfer decreases (due to the conical nature of the Erlenmeyer flask), resulting in decreased aerobicity within the culture *ceteris paribus*. Working volumes of 10 mL, 20 mL, and 30 mL were tested in 125 mL Erlenmeyer flasks. Cultivation was in MOPS minimal media as described in section 3.5 supplemented with 10 g/L glucose at 37°C for 18h, with pH controlled using calcium carbonate. The titers and yields calculated can be seen in Figure 37.

Though there is not much difference between 30 and 20 mL working volumes, there is a slight trend of increased methyl ketone concentration. This can be attributed to the decreased growth rates of the cells when moving towards more anaerobic conditions - in general, rate of product formation can be correlated to specific growth rate.[100] However, in agreement with the hypothesis, the yield of total methyl ketones from glucose slowly increases as aerobicity decreases, inferring that the conversion is more efficient as conditions move towards a microaerobic state. The foreseeable disadvantage

to moving towards microaerobic conditions is that various fermentative pathways that compete with the fatty acid synthesis pathway for carbon become active as conditions move towards decreased oxygen availability. This knowledge has guided subsequent genetic modifications and a more rigorous control of aerobicity that will be detailed in later sections.

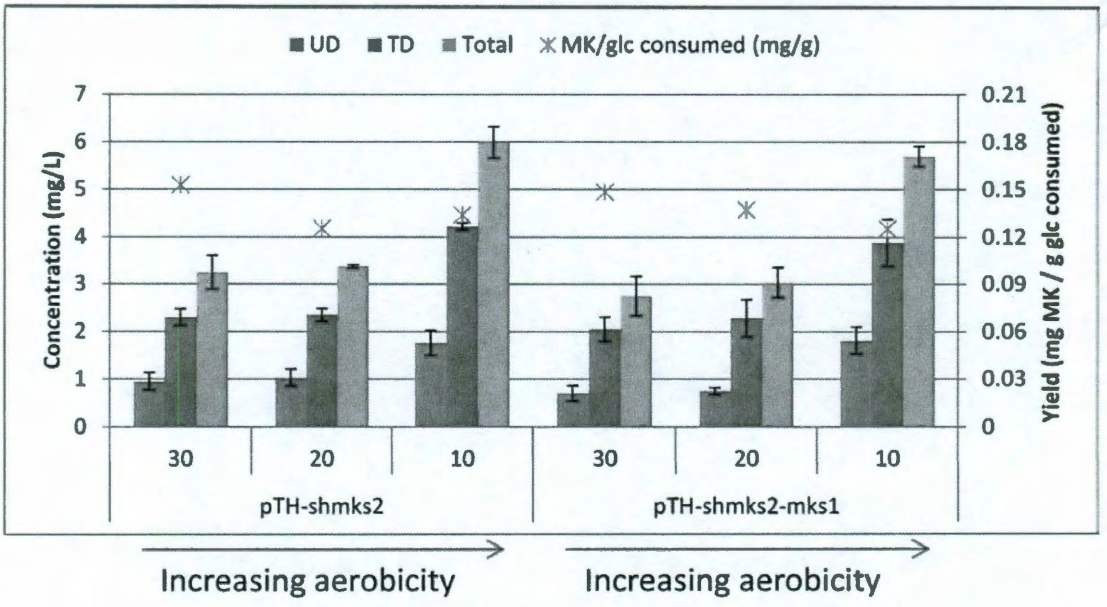



**Figure 37. Effect of aerobicity on methyl ketone titer and yield.**

"UD" denotes 2-undecanone. "TD" denotes 2-tridecanone. "pTH" denotes pTricHis2A. Numbers on the x-axis denote culture working volumes.

## 5.4. ENGINEERED STRAINS FOR THE CONVERSION OF GLUCOSE TO METHYL KETONES UNDER MICROAEROBIC CONDITIONS

Moving towards microaerobic conditions has been found to increase yield of methyl ketones from glucose under the wild-type strain background. However, under anaerobic and microaerobic conditions, *E. coli* begins to convert glucose into a mixture of metabolic products primarily of acetate and formate, with smaller amounts of lactate,

succinate, and ethanol.[99, 101] These pathways are shown in Figure 38. This is problematic as these fermentative pathways consume acetyl-CoA (the building block of fatty acids) and pyruvate (precursor to acetyl-CoA), which may ultimately divert carbon away from methyl ketones. Thus, it was hypothesized that knocking out some or all of these pathways would result in increased yields.

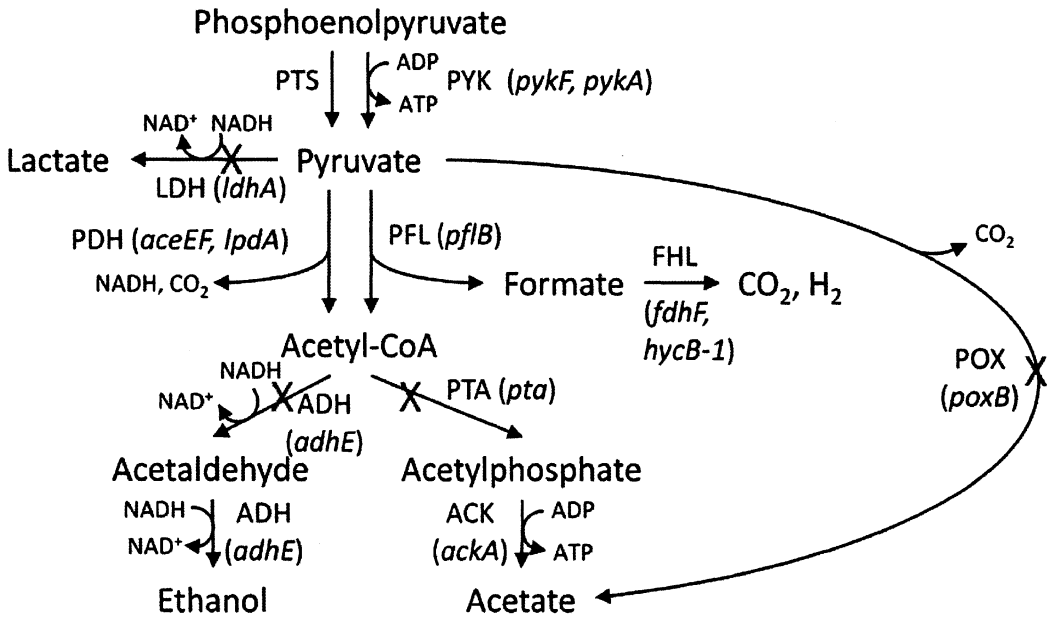



**Figure 38. Mixed acid fermentative pathways and knockouts**

The plasmid pTrcHis2A-shmks2-mks1 was transformed into several mutants that were constructed for this study, which are listed in Table 6 for convenience. The mutants were combinations of the 4 possible pathway knockouts. Cultivation was done in MOPS minimal media supplemented with 10 g/L glucose as described in section 3.5. Shake flasks (125 mL total volume) were prepared with 10, 20, 30, 40, and 75 mL working volumes of media, inoculated from starter cultures to 0.05, and incubated for different

durations depending on the working volume (18h for 10, 20 mL working volumes, 22h for 30, 40, and 75 mL working volumes). Different durations were used due to the lowered growth rates as anaerobic conditions were approached. The lowered growth rates forced higher incubation durations in order to harvest the cells when less than half the glucose remained.

**Table 6. List of MG1655 mutants constructed**

| |
|---|
| ΔadhE |
| ΔldhA |
| ΔpoxB |
| Δpta |
| ΔadhE Δpta |
| ΔldhA ΔpoxB |
| ΔldhA ΔpoxB Δpta |
| ΔadhE ΔldhA ΔpoxB Δpta |

The effects of these knockouts and also the effect of aerobicity for each knockout combination were evaluated by measuring yield and methyl ketone concentration. Comparison with the knockout strains and the wild-type background strain can be done by comparing the data in Figure 37 and Figure 38, as the data in Figure 37 was obtained through the same cultivation conditions as those used for the 10, 20, and 30 mL working volume cultures in Figure 38. The full HPLC profiles (measurement of major acids, pyruvate, glucose remaining) and OD/pH values can be viewed in the appendix A.

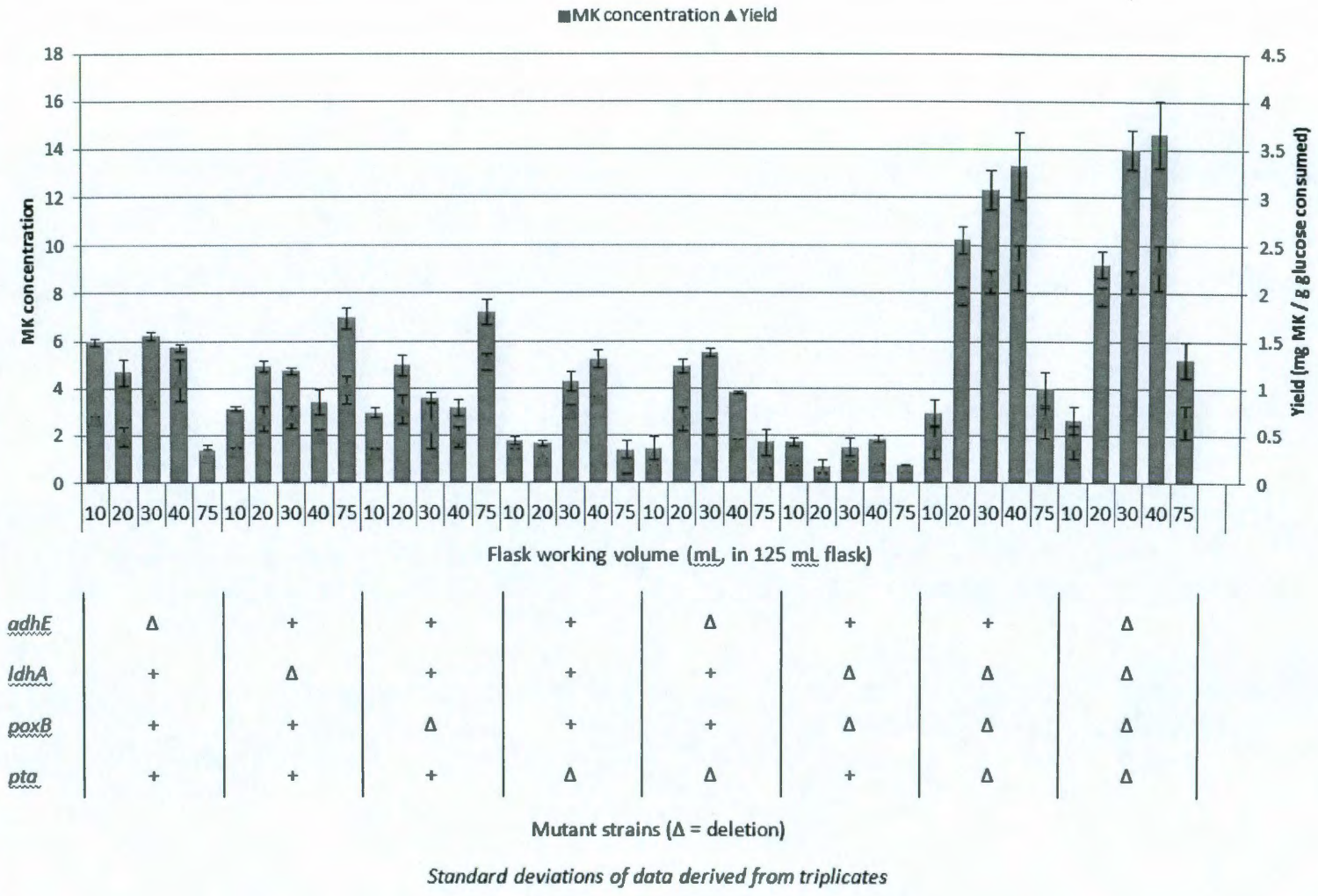



**Figure 39. Effect of aerobicity and various background knockouts on methyl ketone titer and yield.**

The data shows quite clearly that highest yield and titers are attained through strain backgrounds with almost all the fermentative pathways knocked out (ΔldhA ΔpoxB Δpta and ΔadhE ΔldhA ΔpoxB Δpta) as was hypothesized. Almost all the different knockout strains performed slightly better in terms of yield over the wild-type background in varying degrees, which was also expected. Curiously, the double-mutant ΔldhA ΔpoxB gave low titers and yields compared to its respective single-mutants (ΔldhA-only and ΔpoxB-only), with the yield values similar to the non-engineered wild-type. With growth and glucose consumption similar to the other strains, this was unexpected, and HPLC analysis revealed build-up of pyruvate pools much more so than in the other mutant strains. This is partially contrary to the hypothesis that knocking out the competing pathways for pyruvate, thus increasing pools of pyruvate, would lead to increased flux down to acetyl-CoA and ultimately to methyl ketones. However, an additional *pta* deletion relieved this build-up. It is unclear as to why this had occurred, though it remains that the quadruple and triple mutants performed the best in terms of both concentration and yield, with the former being slightly better in these metrics than the latter. This makes sense, as the rationale behind the knockouts was to redirect carbon flux from fermentative pathways that became active in the presence of lesser oxygen. As other metabolite production is increasingly stopped, it is found that there is at least a higher proportion of flux traveling through the fatty acid synthesis pathways, resulting in a subsequent higher increase in both methyl ketone yield and concentration.

## 5.5. STRAIN CHARACTERIZATION UNDER TIGHTLY CONTROLLED OPERATING CONDITIONS

The best performing strain from the shake flask experiments detailed in the previous section (section 5.5) was cultivated in a 500 mL bioreactor vessel in which pH and dissolved oxygen could be tightly controlled. In order to quantify the specific dissolved oxygen (DO) percentage that optimizes yield and titer and to get a better grasp on how these metrics change with specific changes in DO, several bioreactor fermentations were conducted with different DO setpoints. MOPS minimal media as prepared in section 3.5 was supplemented with 20 g/L glucose. The pH was set at 7, temperature at 37°C, and cultures were harvested at 12h. Figure 40 shows how yield and titer change over a range of dissolved oxygen values between 2% and 20%. A table with these values, OD, and HPLC data is included in appendix A.

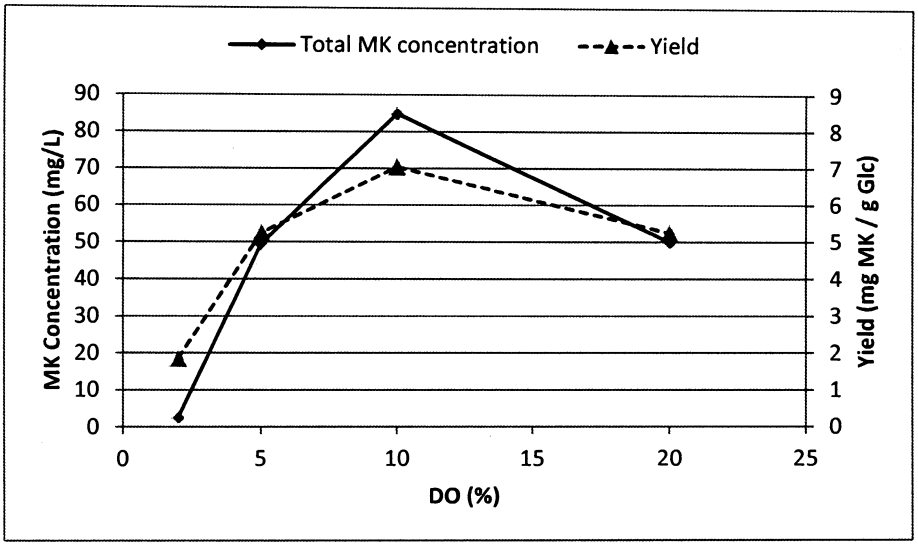



**Figure 40. Methyl ketone concentration and yield as a function of DO (%)**

An unexpected outcome was a marked increase in both the concentration and yield of methyl ketones over those obtained from shake-flasks. There are several parameters that changed when moving from shake-flasks to the bioreactors which may be responsible for this drastic improvement of methyl ketone production. The first is more consistent oxygen availability. One study, which measured dissolved oxygen content over the course of *E. coli* and yeast fermentations in 250-mL shake-flasks with respect to both culture volume and shaker RPM, showed high variability of DO during the fermentations.[102] Specifically, the authors encountered the onset of oxygen limitation (DO measured at ~0%) for both 50 mL and 75 mL culture volumes and at 150, 250, and 350 RPM shaker speeds. During late exponential and early stationary phase, the dissolved oxygen content rapidly increased back to almost 100% saturation. Increases in RPM and culture volumes increased the onset and duration of the oxygen-limiting conditions. Though these results cannot be perfectly extrapolated to explain the conditions in 125-mL shake-flasks with 10 mL cultures, it gives some insight into the conditions of the 30 and 40 mL cultures due to the similarities of the culture volume to flask volume ratios used in the aforementioned study. Conversely, in the bioreactor, DO was monitored and kept steady at specified set-points above 0% saturation, while keeping DO at these setpoints throughout all phases of growth including late exponential and stationary phases instead of allowing it to rise back to high levels of saturation. In other words, the increased availability of oxygen in the media, however low the saturation level, may have greatly improved the production of methyl ketones. Secondly, in the shake-flasks, there may have been losses of methyl ketones through evaporation, as methyl ketones have been

detected in the headspace of methyl ketone-producing *E. coli.* [20] In the bioreactor setup, a condenser is placed between the headspace and the headspace outlet to minimize losses of volatile compounds to the outside. However, according to studies reporting the vapor pressure of various methyl ketones, 2-tridecanone and 2-undecanone are not significantly volatile.[103] The vapor pressures of both compounds are under 1 mm Hg at 25°C - in comparison, at this temperature, water has a vapor pressure of 23.8 mm Hg, ethanol of 55 mm Hg, and acetate of 17 mm Hg. Thirdly, it is widely known that there are optimal ranges of pH for maximum activity/stability of enzymes. In the shake-flask cultures, even when supplemented with calcium carbonate to control pH, the pH would decrease in some cases to 5.0-6.0. In the bioreactor, pH was maintained at a constant 7.0 throughout cultivation. It may be that the enzymes SHMKS2 and MKS1 exhibit much higher activity at the neutral pH of 7, which would explain the rise in MK titer and yield.

Nevertheless, judging from strain performance by the 12h mark, a maximum for yield and concentration is attained when the culture conditions are set at DO 10%. To further characterize the strain under these optimal growth conditions, a fermentation was run using DO set at 10% and glucose supplemented at 50 g/L, with sampling every 6h starting from 12h to 48h after inoculation. The results of the time-course experiment is show in Figure 41, with growth (OD) and glucose consumption in Figure 42.

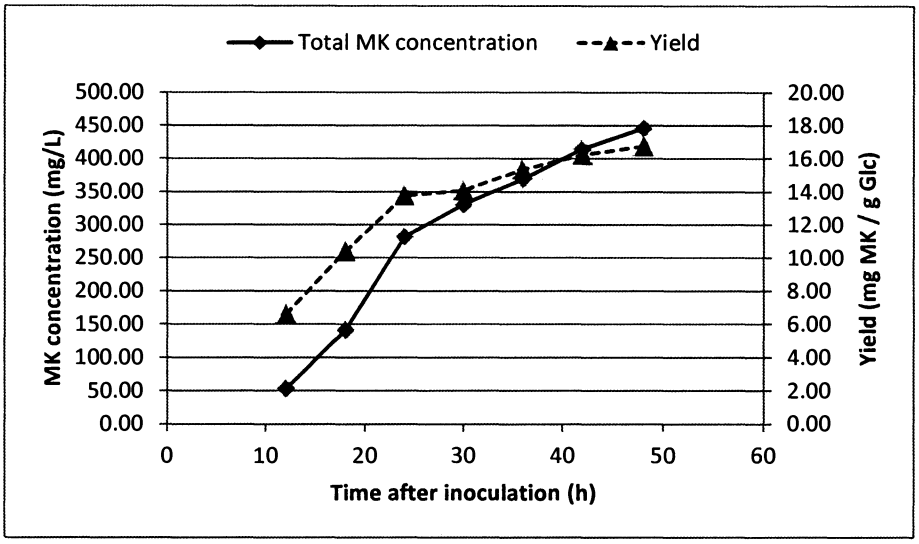



**Figure 41. Methyl ketone (MK) concentration and yield under optimal growth conditions (DO 10%).**

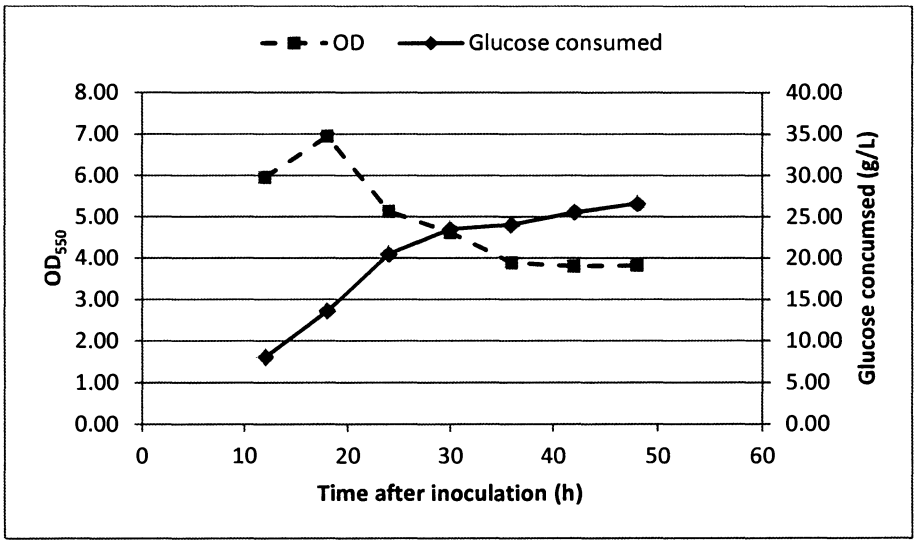



**Figure 42. Glucose consumption and OD under optimal growth conditions (DO 10%).**

The data shows methyl ketone titers approaching 450 mg/L by 48 hours, with yield leveling off at 17 mg MK / g glucose. Both metrics are the highest recorded for any of the strains thus far. Growth entered late exponential to stationary phase by 12 h, at

which point the rate of glucose consumption steadily decreased as expected in concert with the growth profile. If all the glucose in the media were to have been consumed and the yield were to remain constant, the culture would have produced close to 850 mg/L of methyl ketones.

# 6. CONCLUSION

Platform chemicals are a handful of compounds which act as precursors for the synthesis of thousands of high-volume products we use everyday such as plastics, agrochemicals, chemical solvents, etc. Currently, platform chemicals are derived from fossil carbon, which is widely regarded as a non-renewable resource. The long-term future of the chemical industry is thus predicated on the conversion of a renewable resource such as biomass into these platform chemicals, a process which can be accomplished through the metabolic engineering of micro-organisms. In this study, *E. coli* was engineered to produce methyl ketones, a highly reduced platform chemical that has never been synthesized in *E. coli* according to existing literature, at almost 450 mg/L concentrations. Through establishment of optimal operating conditions, and executing on metabolic engineering strategies hypothesized to increase the efficiency of the bioconversion under the established optimal conditions, methyl ketone concentrations were increased almost 75-fold from confirmatory levels of ~6 mg/L to relatively significant levels of ~450 mg/L.

In the future, the optimal operating conditions and genetic background arrived at in this study could act as a starting point for the engineering of new strains that involve grafting other novel pathways that utilize polyketide intermediates. On the other hand,

there are many other modifications proposed to increasing flux through the fatty acid synthesis pathway that do not involve gene deletion such as over-expression of acetyl-CoA carboxylase, over-expression of malic enzyme to increase NADPH pools, and a few others. These modifications have a high chance of success as they have already been published to be viable strategies for increasing fatty acid production.

The improved understanding of the behavior of the polyketide synthesis pathway through evaluation of the effects of various genetic modifications may, in the future, coalesce into well-defined and pertinent metabolic engineering strategies that can be broadly applied towards the efficient bioconversion of glucose to other targets as shown in Figure 43, since they all draw from fatty acid and polyketide metabolism.

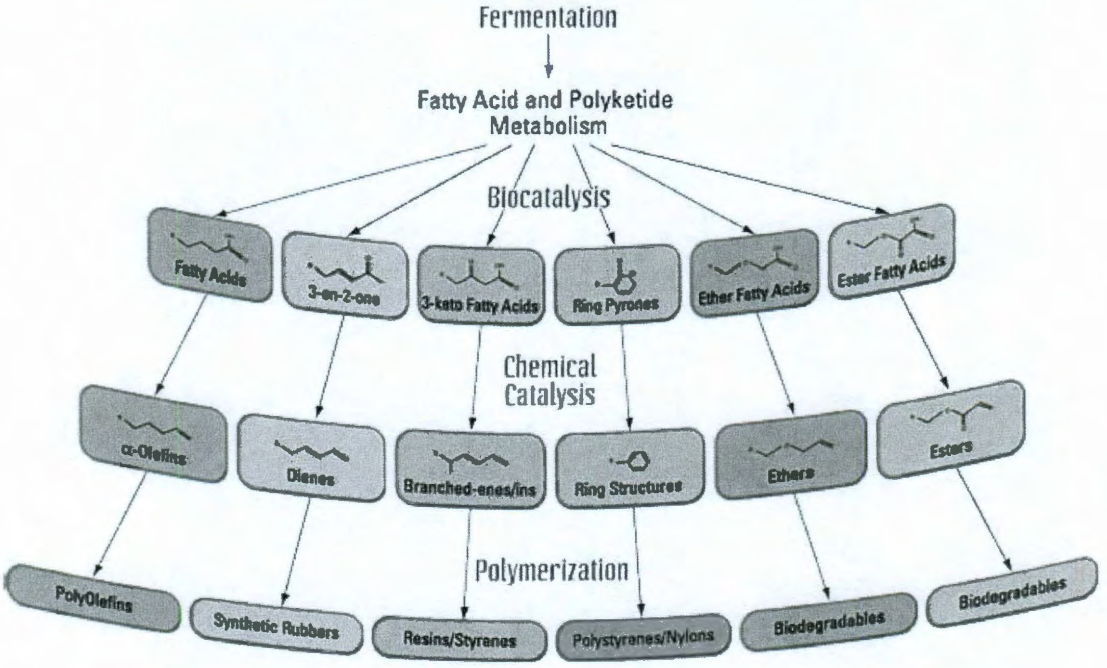



**Figure 43. An array of current and future target molecules that can be derived from polyketide metabolism**

# BIBLIOGRAPHY


1. Doman, L.E. (Department of Energy, Washington DC; 2009).
2. Chenier, P.J. Survey in industrial chemistry, Edn. 3rd. (Kluwer Academic/Plenum Publishers, New York; 2002).
3. Heaton, A., Edn. 3rd (Blackie Academic & Professional, Glasgow, UK; 1991).
4. *McCoy, M.* et al. Fact and figures of the chemical industry. *Chemical & Engineering News* **84**, 35 (2007).
5. OECD Economic Outlook: December No. 76 Volume 2004 Issue 2 Complete Edition - ISBN 9264007768. *OECD Economic Outlook* **2004**, 119-139 (2005).
6. Nikolau, B.J., Perera, M., Brachova, L. & Shanks, B. Platform biochemicals for a biorenewable chemical industry. *Plant Journal* **54**, 536-545 (2008).
7. Corma, A., Iborra, S. & Velty, A. Chemical routes for the transformation of biomass into chemicals. *Chemical Reviews* **107**, 2411-2502 (2007).
8. Werpy, T. & Peterson, G. (US Department of Energy, Washington DC; 2004).
9. Kamm, B. & Kamm, M. Biorefineries - Multi product processes. *White Biotechnology* **105**, 175-204 (2007).
10. Khullar, E., Sall, E.D., Rausch, K.D., Tumbleson, M.E. & Singh, V. Ethanol Production from Modified and Conventional Dry-Grind Processes Using Different Corn Types. *Cereal Chemistry* **86**, 616-622 (2009).
11. McCready, N.S. & Williams, R.C. Utilization of Biofuel Coproducts as Performance Enhancers in Asphalt Binder. *Transportation Research Record*, 8-14 (2008).
12. Langeveld, J.W.A., Dixon, J. & Jaworski, J.F. Development Perspectives Of The Biobased Economy: A Review. *Crop Science* **50**, S142-S151 (2010).
13. Engin, A., Haluk, H. & Gurkan, K. Production of lactic acid esters catalyzed by heteropoly acid supported over ion-exchange resins. *Green Chemistry* **5**, 460-466 (2003).
14. Vu, D.T., Lira, C.T., Asthana, N.S., Kolah, A.K. & Miller, D.J. Vapor-liquid equilibria in the systems ethyl lactate plus ethanol and ethyl lactate plus water. *Journal of Chemical and Engineering Data* **51**, 1220-1225 (2006).
15. Reisch, M.S. Liveris Tells It Like It Is. *Chemical and Engineering News* **84**, 10-15 (2007).
16. Fujita, Y., Matsuoka, H. & Hirooka, K. Regulation of fatty acid metabolism in bacteria. *Molecular Microbiology* **66**, 829-839 (2007).
17. Magnuson, K., Jackowski, S., Rock, C.O. & Cronan, J.E. REGULATION OF FATTY-ACID BIOSYNTHESIS IN ESCHERICHIA-COLI. *Microbiological Reviews* **57**, 522-542 (1993).
18. White, S.W., Zheng, J., Zhang, Y.M. & Rock, C.O. The structural biology of type II fatty acid biosynthesis. *Annual Review of Biochemistry* **74**, 791-831 (2005).
19. Carole, T.M., Pellegrino, J. & Paster, M.D. Opportunities in the industrial biobased products industry. *Applied Biochemistry and Biotechnology* **113**, 871-885 (2004).
20. Fridman, E. et al. Metabolic, genomic, and biochemical analyses of glandular trichomes from the wild tomato species Lycopersicon hirsutum identify a key enzyme in the biosynthesis of methylketones. *Plant Cell* **17**, 1252-1267 (2005).
21. Ben-Israel, I. et al. Multiple Biochemical and Morphological Factors Underlie the Production of Methylketones in Tomato Trichomes. *Plant Physiology* **151**, 1952-1964 (2009).
22. Tseng, G., Sargent, G. & Meadows, J. KIRK-OTHMER ENCYCLOPEDIA OF CHEMICAL-TECHNOLOGY, 3RD EDITION - CD-ROM VERSION. *Nature* **333**, 607-608 (1988).

23. Russell, C.A. The age of polymers and petrochemicals (industrial organic chemistry from 1914). *Chemistry Society and Environment*, 239-269 (2000).
24. Xie, D.M. et al. Microbial synthesis of triacetic acid lactone. *Biotechnology and Bioengineering* **93**, 727-736 (2006).
25. Liao, J.C., Hou, S.Y. & Chao, Y.P. Pathway analysis, engineering, and physiological considerations for redirecting central metabolism. *Biotechnology and Bioengineering* **52**, 129-140 (1996).
26. Carlsen, M., Nielsen, J. & Villadsen, J. Growth and alpha-amylase production by Aspergillus oryzae during continuous cultivations. *Journal of Biotechnology* **45**, 81-93 (1996).
27. Morkeberg, R., Carlsen, M. & Nielsen, J. INDUCTION AND REPRESSION OF ALPHA-AMYLASE PRODUCTION IN BATCH AND CONTINUOUS CULTURES OF ASPERGILLUS-ORYZAE. *Microbiology-Uk* **141**, 2449-2454 (1995).
28. Emmerling, M., Bailey, J.E. & Sauer, U. Glucose catabolism of Escherichia coli strains with increased activity and altered regulation of key glycolytic enzymes. *Metab Eng* **1**, 117-127 (1999).
29. Cronan, J.E. & Waldrop, G.L. Multi-subunit acetyl-CoA carboxylases. *Progress in Lipid Research* **41**, 407-435 (2002).
30. Davis, M.S., Solbiati, J. & Cronan, J.E. Overproduction of acetyl-CoA carboxylase activity increases the rate of fatty acid biosynthesis in Escherichia coli. *Journal of Biological Chemistry* **275**, 28593-28598 (2000).
31. Lennen, R.M., Braden, D.J., West, R.M., Dumesic, J.A. & Pfleger, B.F. A Process for Microbial Hydrocarbon Synthesis: Overproduction of Fatty Acids in Escherichia coli and Catalytic Conversion to Alkanes. *Biotechnology and Bioengineering* **106**, 193-202 (2010).
32. Lu, X.F., Vora, H. & Khosla, C. Overproduction of free fatty acids in E. coli: Implications for biodiesel production. *Metabolic Engineering* **10**, 333-339 (2008).
33. Sheldon, R.A. Chirotechnology. (Marcel Dekker, New York; 1993).
34. Watson, J.D. & Crick, F.H.C. MOLECULAR STRUCTURE OF NUCLEIC ACIDS - A STRUCTURE FOR DEOXYRIBOSE NUCLEIC ACID. *Nature* **171**, 737-738 (1953).
35. Cohen, S.N., Chang, A.C.Y., Boyer, H.W. & Helling, R.B. CONSTRUCTION OF BIOLOGICALLY FUNCTIONAL BACTERIAL PLASMIDS IN-VITRO. *Proceedings of the National Academy of Sciences of the United States of America* **70**, 3240-3244 (1973).
36. . (ed. B.R.D.T.A. Committee) (U.S. Department of Energy, 2002).
37. Leeper, S.A. & Andrews, G.F. A CRITICAL-REVIEW AND EVALUATION OF BIOPRODUCTION OF ORGANIC-CHEMICALS. *Applied Biochemistry and Biotechnology* **28-9**, 499-511 (1991).
38. Jung, Y.K., Kim, T.Y., Park, S.J. & Lee, S.Y. Metabolic Engineering of Escherichia coli for the Production of Polylactic Acid and Its Copolymers. *Biotechnology and Bioengineering* **105**, 161-171 (2010).
39. Singh, S.K., Ahmed, S.U. & Pandey, A. Metabolic engineering approaches for lactic acid production. *Process Biochemistry* **41**, 991-1000 (2006).
40. Lee, S.J. et al. Metabolic engineering of Escherichia coli for enhanced production of succinic acid, based on genome comparison and in silico gene knockout simulation. *Applied and Environmental Microbiology* **71**, 7880-7887 (2005).
41. Bunch, P.K., MatJan, F., Lee, N. & Clark, D.P. The IdhA gene encoding the fermentative lactate dehydrogenase of Escherichia coli. *Microbiology-Uk* **143**, 187-195 (1997).
42. Millard, C.S., Chao, Y.P., Liao, J.C. & Donnelly, M.I. Enhanced production of succinic acid by overexpression of phosphoenolpyruvate carboxylase in Escherichia coli. *Applied and Environmental Microbiology* **62**, 1808-1810 (1996).

43. Hong, S.H. & Lee, S.Y. Metabolic flux analysis for succinic acid production by recombinant Escherichia coli with amplified malic enzyme activity. *Biotechnology and Bioengineering* **74**, 89-95 (2001).
44. Gokarn, R.R. et al. The physiological effects and metabolic alterations caused by the expression of Rhizobium etli pyruvate carboxylase in Escherichia coli. *Applied Microbiology and Biotechnology* **56**, 188-195 (2001).
45. Williams, C.G. On the Constitution of the Essential Oil of Rue. [Abstract]. *Proceedings of the Royal Society of London* **9**, 167-169 (1857).
46. Watts, F. On the essential oil of lime leaves (citrus limetta). *Journal of the Chemical Society, Transactions* **49**, 316-317 (1886).
47. Thoms, H. Die Zusammensetzung des aetherischen Rautenoeles. *Ber. Deut. Pharm. Ges.* **11**, 3-19 (1901).
48. Houben, J. Ueber das deutsche Rautenoel und die Umlagerung des Methylnonylketoxims. (1902).
49. Thoms, H. Ueber die Wertbestimmung des Nelkenoeles. *Arch. Pharm.* **241**, 592-603 (1903).
50. Wehmer, C. Die Pflanzenstoffe, Vol. 1. (Edward Brothers, Inc., Ann Arbor, Mich.; 1950).
51. Morgan, M.E. & Anderson, E.O. THE NEUTRAL CARBONYL COMPOUNDS IN BLUE-MOLD TYPE CHEESE. *Journal of Dairy Science* **39**, 253-260 (1956).
52. Petersen, D.J. & Bennett, G.N. PURIFICATION OF ACETOACETATE DECARBOXYLASE FROM CLOSTRIDIUM-ACETOBUTYLICUM ATCC-824 AND CLONING OF THE ACETOACETATE DECARBOXYLASE GENE IN ESCHERICHIA-COLI. *Applied and Environmental Microbiology* **56**, 3491-3498 (1990).
53. Ravagnani, A. et al. Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. *Molecular Microbiology* **37**, 1172-1185 (2000).
54. Schulz, S. & Dickschat, J.S. Bacterial volatiles: the smell of small organisms. *Natural Product Reports* **24**, 814-842 (2007).
55. Patel, R.N., Hou, C.T., Laskin, A.I., Derelanko, P. & Felix, A. MICROBIAL-PRODUCTION OF METHYL KETONES - PURIFICATION AND PROPERTIES OF A SECONDARY ALCOHOL-DEHYDROGENASE FROM YEAST. *European Journal of Biochemistry* **101**, 401-406 (1979).
56. Kennedy, G.G. Tomato, Pests, parasitoids, and predators: Tritrophic interactions involving the genus Lycopersicon. *Annual Review of Entomology* **48**, 51-72 (2003).
57. Williams, W.G., Kennedy, G.G., Yamamoto, R.T., Thacker, J.D. & Bordner, J. 2-TRIDECANONE - NATURALLY-OCCURRING INSECTICIDE FROM THE WILD TOMATO LYCOPERSICON-HIRSUTUM F-GLABRATUM. *Science* **207**, 888-889 (1980).
58. Hotelier, T. et al. ESTHER, the database of the alpha/beta-hydrolase fold superfamily of proteins. *Nucleic Acids Research* **32**, D145-D147 (2004).
59. Dillon, S.C. & Bateman, A. The Hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. *Bmc Bioinformatics* **5** (2004).
60. Shearer, D.A. & Boch, R. 2-HEPTANONE IN MANDIBULAR GLAND SECRETION OF HONEY-BEE. *Nature* **206**, 530-& (1965).
61. Moser, J.C., Brownlee, R.C. & Silverst.R ALARM PHEROMONES OF ANT ATTA TEXANA. *Journal of Insect Physiology* **14**, 529-& (1968).
62. Bernardi, R. et al. On the components of secretion of mandibular glands of the ant Lasius (Dendrolasius) fuligtnosus. *Tetrahedron Lett* **4**, 3893-3896 (1967).
63. Regnier, F.E. & Wilson, E.O. The alarm-defence system of the ant Acanthomyops claviger. *Journal of Insect Physiology* **14**, 955-970 (1968).

64. Gehrig, R.F. & Knight, S.G. FORMATION OF 2-HEPTANONE FROM CAPRYLIC ACID BY SPORES OF VARIOUS FILAMENTOUS FUNGI. *Nature* **192**, 1185-& (1961).
65. Gehrig, R.F. & Knight, S.G. FORMATION OF KETONES FROM FATTY ACIDS BY SPORES OF PENICILLIUM-ROQUEFORTI. *Nature* **182**, 1237-1237 (1958).
66. Kinsella, J.E. & Hwang, D.H. ENZYMES OF PENICILLIUM ROQUEFORTI INVOLVED IN BIOSYNTHESIS OF CHEESE FLAVOR. *Crc Critical Reviews in Food Science and Nutrition* **8**, 191-228 (1976).
67. Dartey, C.K. & Kinsella, J.E. OXIDATION OF SODIUM U-C-14 PALMITATE INTO CARBONYL-COMPOUNDS BY PENICILLIUM-ROQUEFORTI SPORES. *Journal of Agricultural and Food Chemistry* **21**, 721-726 (1973).
68. Britt, A.J., Bruce, N.C. & Lowe, C.R. Purification and characterisation of an NAD(+)-dependent secondary alcohol dehydrogenase from Pseudomonas maltophilia MB11L. *FEMS Microbiol Lett* **72**, 49-55 (1992).
69. Hou, C.T., Patel, R., Laskin, A.I., Barnabe, N. & Marczak, I. MICROBIAL OXIDATION OF GASEOUS HYDROCARBONS - PRODUCTION OF METHYL KETONES FROM THEIR CORRESPONDING SECONDARY ALCOHOLS BY METHANE-GROWN AND METHANOL-GROWN MICROBES. *Applied and Environmental Microbiology* **38**, 135-142 (1979).
70. Hou, C.T., Patel, R., Laskin, A.I., Barnabe, N. & Barist, I. PRODUCTION OF METHYL KETONES FROM SECONDARY ALCOHOLS BY CELL-SUSPENSIONS OF C-2 TO C-4 N-ALKANE-GROWN BACTERIA. *Applied and Environmental Microbiology* **46**, 178-184 (1983).
71. Hou, C.T., Patel, R.N., Laskin, A.I., Barist, I. & Barnabe, N. THERMOSTABLE NAD-LINKED SECONDARY ALCOHOL-DEHYDROGENASE FROM PROPANE-GROWN PSEUDOMONAS-FLUORESCENS NRRL-B-1244. *Applied and Environmental Microbiology* **46**, 98-105 (1983).
72. Cornelis, P. Pseudomonas: Genomics and Molecular Biology, Edn. 1st. (Caister Academic Press, 2008).
73. Kirschner, A., Altenbuchner, J. & Bornscheuer, U.T. Design of a secondary alcohol degradation pathway from Pseudomonas fluorescens DSM 50106 in an engineered Escherichia coli. *Applied Microbiology and Biotechnology* **75**, 1095-1101 (2007).
74. Hockelmann, C. & Juttner, F. Volatile organic compound (VOC) analysis and sources of limonene, cyclohexanone and straight chain aldehydes in axenic cultures of Calothrix and Plectonema. *Water Science and Technology* **49**, 47-54 (2004).
75. Bruce, A., Verrall, S., Hackett, C.A. & Wheatley, R.E. Identification of volatile organic compounds (VOCs) from bacteria and yeast causing growth inhibition of sapstain fungi. *Holzforschung* **58**, 193-198 (2004).
76. Fadda, S., Leroy-Setrin, S. & Talon, R. Preliminary characterization of beta-decarboxylase activities in Staphylococcus carnosus 833, a strain used in sausage fermentation. *Fems Microbiology Letters* **228**, 143-149 (2003).
77. Lehninger, A., Nelson, D. & Cox, M. Lehninger Principles of Biochemistry. (W. H. Freeman, 2008).
78. Rock, C.O. & Jackowski, S. REGULATION OF PHOSPHOLIPID-SYNTHESIS IN ESCHERICHIA-COLI - COMPOSITION OF THE ACYL-ACYL CARRIER PROTEIN POOL INVIVO. *Journal of Biological Chemistry* **257**, 759-765 (1982).
79. Polakis, S.E., Guchhait, R.B. & Lane, M.D. STRINGENT CONTROL OF FATTY-ACID SYNTHESIS IN ESCHERICHIA-COLI - POSSIBLE REGULATION OF ACETYL COENZYME-A CARBOXYLASE BY PPGPP. *Journal of Biological Chemistry* **248**, 7957-7966 (1973).

80. Li, S.J. & Cronan, J.E. GROWTH-RATE REGULATION OF ESCHERICHIA-COLI ACETYL COENZYME-A CARBOXYLASE, WHICH CATALYZES THE 1ST COMMITTED STEP OF LIPID BIOSYNTHESIS. *Journal of Bacteriology* **175**, 332-340 (1993).
81. Heath, R.J. & Rock, C.O. Regulation of fatty acid elongation and initiation by acyl acyl carrier protein in Escherichia coli. *Journal of Biological Chemistry* **271**, 1833-1836 (1996).
82. Subrahmanyam, S. & Cronan, J.E. Overproduction of a functional fatty acid biosynthetic enzyme blocks fatty acid synthesis in Escherichia coli. *Journal of Bacteriology* **180**, 4596-4602 (1998).
83. Podkovyrov, S.M. & Larson, T.J. Identification of promoter and stringent regulation of transcription of the fabH, fabD and fabG genes encoding fatty acid biosynthetic enzymes of Escherichia coli. *Nucleic Acids Research* **24**, 1747-1752 (1996).
84. Heath, R.J. & Rock, C.O. Roles of the FabA and FabZ beta-hydroxyacyl-acyl carrier protein dehydratases in Escherichia coli fatty acid biosynthesis. *Journal of Biological Chemistry* **271**, 27795-27801 (1996).
85. Clark, D.P., Demendoza, D., Polacco, M.L. & Cronan, J.E. BETA-HYDROXYDECANOYL THIO ESTER DEHYDRASE DOES NOT CATALYZE A RATE-LIMITING STEP IN ESCHERICHIA-COLI UNSATURATED FATTY-ACID SYNTHESIS. *Biochemistry* **22**, 5897-5902 (1983).
86. Voelker, T.A. & Davies, H.M. ALTERATION OF THE SPECIFICITY AND REGULATION OF FATTY-ACID SYNTHESIS OF ESCHERICHIA-COLI BY EXPRESSION OF A PLANT MEDIUM-CHAIN ACYL-ACYL CARRIER PROTEIN THIOESTERASE. *Journal of Bacteriology* **176**, 7320-7327 (1994).
87. Wynn, J.P. & Ratledge, C. Malic enzyme is a major source of NADPH for lipid accumulation by Aspergillus nidulans. *Microbiology-Uk* **143**, 253-257 (1997).
88. Bronlund, J.E. Separating the Meat from the Soup: Bioseparations Science and Engineering, R.G. Harrison, P. Todd, S.R. Rudge and D.P. Petrides, Oxford University Press, Oxford, 2003, pp. xix+406, ISBN 0-19-512340-9. *Biotechnology Advances* **21**, 211-212 (2003).
89. Bligh, E.G. & Dyer, W.J. A RAPID METHOD OF TOTAL LIPID EXTRACTION AND PURIFICATION. *Canadian Journal of Biochemistry and Physiology* **37**, 911-917 (1959).
90. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I-1-I-2 regulatory elements. *Nucleic Acids Research* **25**, 1203-1210 (1997).
91. Bachler, C., Schneider, P., Bahler, P., Lustig, A. & Erni, B. Escherichia coli dihydroxyacetone kinase controls gene expression by binding to transcription factor DhaR. *Embo Journal* **24**, 283-293 (2005).
92. Yazdani, S.S. & Gonzalez, R. Engineering Escherichia coli for the efficient conversion of glycerol to ethanol and co-products. *Metabolic Engineering* **10**, 340-351 (2008).
93. Hayes, W.S. & Borodovsky, M. Deriving ribosomal binding site (RBS) statistical models from unannotated DNA sequences and the use of the RBS model for N-terminal prediction. *Pac Symp Biocomput*, 279-290 (1998).
94. Kang, Y.S. et al. Systematic mutagenesis of the Escherichia coli genome (vol 186, pg 4921, 2004). *Journal of Bacteriology* **186**, 8548-8548 (2004).
95. Blankschien, M.D., Clomburg, J. & Gonzalez, R. Metabolic engineering of Escherichia coli for the production of succinate from glycerol. *Metabolic Engineering* **12**, 409-419 (2010).
96. Neidhard.Fc, Bloch, P.L. & Smith, D.F. CULTURE MEDIUM FOR ENTEROBACTERIA. *Journal of Bacteriology* **119**, 736-747 (1974).
97. Sargent, T. Current protocols in molecular biology : Edited by , and , Wiley, New York, 1987. $140.00. *Analytical Biochemistry* **171**, 223-224 (1988).

98. Dharmadi, Y. & Gonzalez, R. A better global resolution function and a novel iterative stochastic search method for optimization of high-performance liquid chromatographic separation. *Journal of Chromatography A* **1070**, 89-101 (2005).
99. Durnin, G. et al. Understanding and Harnessing the Microaerobic Metabolism of Glycerol in Escherichia coli. *Biotechnology and Bioengineering* **103**, 148-161 (2009).
100. Omstead, D.R. Computer control of fermentation processes, Edn. 1st Edition. (CRC Press, 1989).
101. Clark, D.P. THE FERMENTATION PATHWAYS OF ESCHERICHIA-COLI. *Fems Microbiology Reviews* **63**, 223-234 (1989).
102. Tolosa, L., Kostov, Y., Harms, P. & Rao, G. Noninvasive measurement of dissolved oxygen in shake flasks. *Biotechnology and Bioengineering* **80**, 594-597 (2002).
103. Meyer, E.F. & Wagner, R.E. COHESIVE ENERGIES IN POLAR ORGANIC LIQUIDS. *Journal of Physical Chemistry* **70**, 3162-& (1966).
104. Onaca, C., Kieninger, M., Engesser, K.-H. & Altenbuchner, J. Degradation of Alkyl Methyl Ketones by Pseudomonas veronii MEK700. *The Journal of Bacteriology* **189**, 3759-3767 (2007).
105. DiRusso, C.C. & Black, P.N. Long-chain fatty acid transport in bacteria and yeast. Paradigms for defining the mechanism underlying this protein-mediated process. *Molecular and Cellular Biochemistry* **192**, 41-52 (1999).

# 7. APPENDIX A

## 7.1. TRANSPORT OF METHYL KETONES

The standard lysis and extraction method used to quantify methyl ketones after cultivation involved extraction from the culture broth which included the lysed contents of the cells. Methyl ketone concentrations in supernatant and sonicated culture broth after 18h cultivation in MOPS minimal media supplemented with 10 g/L glucose is shown in Figure 44. The graph shows clearly that a majority of the methyl ketones remain as intracellular metabolites.

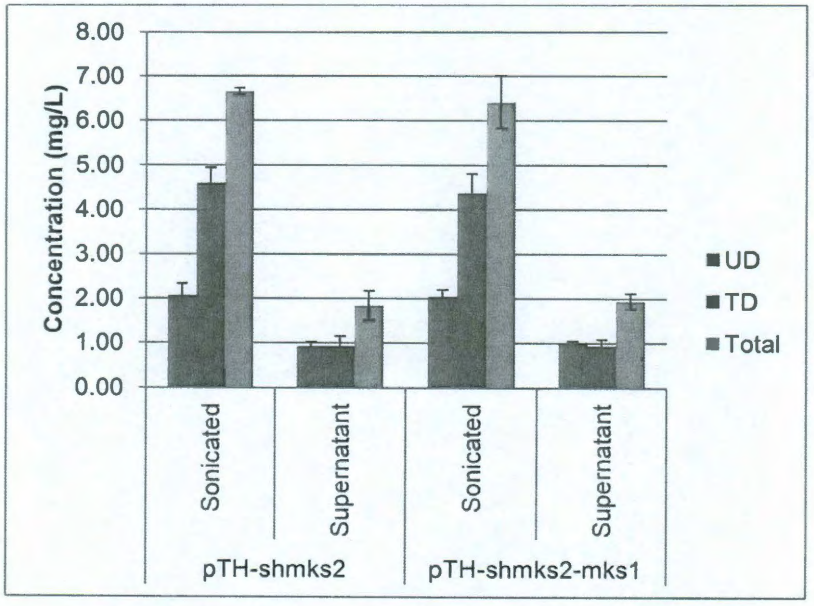



**Figure 44. Methyl ketone production of two strains.**

Note: pTH refers to pTrcHis2A.

Based on this data, methyl ketones are not excreted into the extracellular medium. The identification of natural transporters of methyl ketones could increase the efficiency of the process not only by increasing the ease of extraction by avoiding the need to break down cells, but also increase the efficiency of the reaction itself by pushing it towards

product formation through Le Chatelier's principle. The methyl ketone transporters are currently unknown for both methyl ketone producers as well as pethyl ketone utilizers [104]. However, there is already a native unspecified transporter that is able to transport methyl ketones both into and outside of the cytoplasm, as seen from some studies and the results shown in figures 35 and 36. Several other targets for this include fatty acid transporters (as methyl ketones are fatty-acid like molecules), specifically the outer membrane protein FadL, periplasmic protein Tsp, and the inner membrane protein FadD [105]. Another possible target is a wild tomato plasma membrane intrinsic protein (GenBank accession BI925004), which may be a transporter of methyl ketones [21]. This protein was one of those that were singled out through an aforementioned transcriptome analysis of methyl ketone-producing plants versus non-producers (the list is shown in Table 2, page 34). It was overexpressed 39-fold in plants producing methyl ketones compared to those producing little to no methyl ketones.

## 7.2. HPLC DATA FOR SHAKE FLASK EXPERIMENTS

**Table 7. pH, OD, and HPLC data for knockout strains in shake flask experiments**

| | Working volume (mL) | pH | OD | Glucose consumed (g/L) | Lactate (g/L) | Ethanol (g/L) | Acetate (g/L) | Succinate (g/L) | Formate (g/L) | Pyruvate (g/L) |
|---|---|---|---|---|---|---|---|---|---|---|
| adhE knockout | 10 | 5.2 | 2.16 | 8.89 | 0.407 | 0.000 | 1.276 | 0.113 | 0.184 | 1.788 |
| | 20 | 7.14 | 4.4 | 9.57 | 0.298 | 0.000 | 0.852 | 0.027 | 0.118 | 0.000 |
| | 30 | 5.54 | 1.7 | 7.27 | 0.249 | 0.000 | 1.043 | 0.219 | 0.212 | 1.697 |
| | 40 | 5.66 | 1.24 | 5.27 | 0.250 | 0.000 | 0.970 | 0.195 | 0.287 | 1.507 |
| | 75 | 5.68 | 0.66 | 4.87 | - | - | - | - | - | - |
| ldhA knockout | 10 | 6.31 | 3.8 | 9.38 | 0.399 | 0.000 | 0.113 | 0.062 | 0.037 | 0.576 |
| | 20 | 6.57 | 1.96 | 7.23 | 1.078 | 0.235 | 0.060 | 0.183 | 0.342 | 0.497 |
| | 30 | 5.02 | 1.5 | 6.81 | 0.595 | 0.180 | 0.098 | 0.206 | 0.179 | 0.891 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 40 | 5.23 | 1.24 | 7.05 | 0.733 | 0.221 | 0.055 | 0.253 | 0.223 | 0.991 |
| | 75 | 6.78 | 1.06 | 7.03 | - | - | - | - | - | - |
| poxB knockout | 10 | 6.85 | 4 | 9.45 | 0.745 | 0.000 | 0.370 | 0.047 | 0.000 | 0.831 |
| | 20 | 6.58 | 3 | 6.44 | 3.021 | 0.000 | 0.331 | 0.078 | 0.100 | 0.858 |
| | 30 | 7.02 | 2.1 | 5.93 | 4.479 | 0.000 | 0.589 | 0.153 | 0.065 | 0.000 |
| | 40 | 6.94 | 1.92 | 6.65 | 4.764 | 0.000 | 0.380 | 0.193 | 0.000 | 0.000 |
| | 75 | 6.75 | 1.24 | 5.66 | - | - | - | - | - | - |
| pta knockout | 10 | 6.31 | 4.74 | 5.87 | 0.032 | 0.051 | 2.351 | 0.075 | 0.037 | 0.669 |
| | 20 | 6.57 | 3.88 | 7.30 | 0.000 | 0.314 | 2.144 | 0.087 | 0.416 | 2.207 |
| | 30 | 5.02 | 1.4 | 5.71 | 0.000 | 0.456 | 3.747 | 0.279 | 0.903 | 0.000 |
| | 40 | 5.23 | 1.46 | 5.88 | 0.000 | 0.544 | 3.589 | 0.591 | 1.308 | 0.209 |
| | 75 | 7.02 | 1.88 | 8.07 | - | - | - | - | - | - |
| adhE/pta knockout | 10 | 6.62 | 2.88 | 8.73 | 0.070 | 0.000 | 0.681 | 0.083 | 0.000 | 0.213 |
| | 20 | 5.43 | 2.2 | 7.39 | 0.010 | 0.049 | 0.967 | 0.170 | 0.158 | 1.586 |
| | 30 | 5.5 | 1.5 | 9.31 | 0.012 | 0.156 | 1.047 | 0.150 | 0.251 | 1.752 |
| | 40 | 5.55 | 1.12 | 9.46 | 0.016 | 0.210 | 1.027 | 0.154 | 0.368 | 1.677 |
| | 75 | 5.97 | 0.63 | 8.98 | - | - | - | - | - | - |
| ldhA/poxB knockout | 10 | 6.96 | 3.3 | 9.63 | 0.042 | 0.000 | 0.585 | 0.064 | 0.000 | 0.000 |
| | 20 | 5.46 | 2.14 | 9.10 | 0.000 | 0.063 | 0.788 | 0.170 | 0.139 | 1.600 |
| | 30 | 5.32 | 1.58 | 7.64 | 0.020 | 0.169 | 1.005 | 0.167 | 0.304 | 2.122 |
| | 40 | 5.5 | 1.22 | 9.85 | 0.013 | 0.252 | 0.963 | 0.179 | 0.407 | 2.039 |
| | 75 | 5.69 | 0.63 | 8.70 | - | - | - | - | - | - |
| ldhA/poxB/ pta knockout | 10 | 6.7 | 3.24 | 6.92 | 0.000 | 0.000 | 0.120 | 0.073 | 0.000 | 0.346 |
| | 20 | 5.42 | 1.7 | 5.20 | 0.000 | 0.000 | 0.049 | 0.077 | 0.030 | 1.041 |
| | 30 | 5.71 | 1.52 | 5.77 | 0.000 | 0.000 | 0.048 | 0.073 | 0.040 | 1.269 |
| | 40 | 5.63 | 1.08 | 5.83 | 0.000 | 0.000 | 0.050 | 0.129 | 0.066 | 1.483 |
| | 75 | 5.58 | 0.87 | 6.04 | - | - | - | - | - | - |
| adhE/ldhA/ poxB/pta knockout | 10 | 6.7 | 3.11 | 6.23 | - | - | - | - | - | - |
| | 20 | 7.25 | 2.82 | 4.69 | 0.056 | 0.000 | 0.052 | 0.071 | 0.064 | 0.415 |
| | 30 | 7.07 | 2.1 | 6.57 | 0.033 | 0.000 | 0.043 | 0.153 | 0.225 | 1.356 |
| | 40 | 7.09 | 1.4 | 6.42 | 0.000 | 0.000 | 0.027 | 0.096 | 0.233 | 0.919 |
| | 75 | 6.88 | 1 | 8.00 | - | - | - | - | - | - |

## 7.3. HPLC DATA FOR FERMENTOR EXPERIMENTS

**Table 8. Tabulated data for fermentor experiment varying DO from 2 to 20%.**

| Time (h) | DO | 2-TD concentration (mg/L) | 2-UD concentration (mg/L) | Total MKs (mg/L) | Glucose consumed (g/L) | Yield (mg MK / g Glc) | Yield (mg MK / OD) | OD | Lactate (g/L) | Ethanol (g/L) | Acetate (g/L) | Succinate (g/L) | Formate (g/L) | Pyruvate (g/L) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12 | 50% | 48 | 2 | 50 | - | 5.25 | - | - | - | - | - | - | - | - |
| 12 | 10% | 79.8 | 5.2 | 85.00 | 12.08 | 7.04 | 15.51 | 5.48 | 0.05 | 0.00 | 0.19 | 0.08 | 0.08 | 1.02 |
| 12 | 5% | 46.6 | 2.96 | 49.56 | 9.43 | 5.25 | 9.64 | 5.14 | 0.00 | 0.00 | 0.13 | 0.18 | 0.27 | 1.76 |
| 12 | 2% | 2.54 | ND | 2.54 | 1.36 | 1.86 | 6.05 | 0.42 | 0.00 | 0.00 | 0.00 | 0.04 | 0.00 | 0.16 |
| 24 | 2% | 6.19 | ND | 6.19 | 2.94 | 2.10 | 8.97 | 0.69 | 0.00 | 0.00 | 0.00 | 0.03 | 0.00 | 0.25 |
| 36 | 2% | 20 | ND | 20.00 | 12.43 | 1.61 | 3.29 | 6.08 | 0.18 | 0.00 | 0.24 | 0.09 | 0.00 | 1.06 |
| 48 | 2% | 23 | 1.92 | 24.92 | 18.57 | 1.34 | 2.80 | 8.90 | 0.07 | 0.00 | 0.61 | 0.11 | 0.00 | 3.40 |